$7.00 per volume
Single numbers $4.00

QUARTERLY

Two volumes per year

July, 1936
Volume XV, No. 1

Founded in 1927 by Edward Bradford Titchener and Carl Murchison

THE JOURNAL OF GENERAL PSYCHOLOGY

Experimental, Theoretical, Clinical, and Historical Psychology

JULY, 1936

(Manuscripts are printed in the order of final acceptance)

THE DEPENDENCE OF THE ABSOLUTE VISUAL THRESHOLD ON THE PRESENCE OF AN ADDITIONAL STIMULUS IN THE FIELD OF VISION 3
By B. M. Teplov

RECOVERY OF SENSITIVITY TO PRICK AND TOUCH AFTER PRESSURE BLOCK, WITH A NOTE ON THE SENSITIVITY OF HAIR SPOTS 13
By Gertrude Raffel

LURIA'S MOTOR METHOD AND WORD ASSOCIATION IN THE STUDY OF DECEPTION 23
By Joseph E. Runkel

THE SIGNIFICANCE OF THE FLUCTUATIONS EXPERIENCED IN OBSERVING AMBIGUOUS FIGURES AND IN BINOCULAR RIVALRY 39
By R. Wilfred George

STUDIES OF MOTOR FUNCTION IN SCHIZOPHRENIA: I. SPEED OF TAPPING 63
By David Shakow and Paul E. Huston

(over)

Copyright, 1936, by Clark University
Worcester, Massachusetts

Entered as second-class matter August 3, 1897, at the post-office at Worcester, Mass., under Act of March 3, 1879.

A TECHNIQUE FOR INVESTIGATING RETROACTIVE AND OTHER INHIBITORY EFFECTS IN IMMEDIATE MEMORY 107
By James J. Gibson and Gertrude Raffel

ON THE FUNCTIONAL RELATIONS OF THE SENSE ORGANS TO ONE ANOTHER AND TO THE ORGANISM AS A WHOLE 117
By Walter Börnstein

THE FUSION FREQUENCY IN DIFFERENT AREAS OF THE VISUAL FIELD: I. THE FOVEAL FUSION FREQUENCY 133
By Robert T. Ross

THE TOUCH SPOTS AND THE INTENSITY OF THE STIMULUS 149
By J. P. Guilford and Edna M. Lovewell

THE FUSION FREQUENCY IN DIFFERENT AREAS OF THE VISUAL FIELD: II. THE REGIONAL GRADIENT OF FUSION FREQUENCY 161
By Robert T. Ross

DYNAMIC AND TEMPORAL CONTROL IN MUSIC . . 171
By O. Irving Jacobsen

SHORT ARTICLES AND NOTES 191
Relation of Suggestibility to Other Personality Traits (Marion R. Bartlett) 191
Types and Methods of Learning (Carl N. Rexroad) . . 196
Thirst As an Arbitrary Drive (B. F. Skinner) . . . 205
Sleep, Work, and Food Habits in the Tropics (Richard Wellington Husband) 210
Psychophysics and Mescal Intoxication (Theodore Karowoski) 212

APPARATUS 221
An Inexpensive Complication Device (Rex Madison Collier) 221

BOOKS 225

BOOKS RECEIVED 227

THE DEPENDENCE OF THE ABSOLUTE VISUAL THRESHOLD ON THE PRESENCE OF AN ADDITIONAL STIMULUS IN THE FIELD OF VISION*

From the State Institute for Psychology, Moscow

B. M. Teplov

Problem

The value of the absolute visual threshold depends on the presence of other light stimuli in the field of vision. This dependence has been stated by many authors, but its character proved different under different experimental conditions. The two following groups of facts are the most essential in this respect.

On one hand it may be considered firmly established that the threshold value for several simultaneously acting spatially separate stimuli is less than that of one such stimulus. This phenomenon has been observed by Charpentier (2), Fick (3), Piper (9). Recently it was investigated in detail by Beitel (1) and by Sevrugueena in collaboration with the author of this paper (10). Beitel found that the threshold for two semi-circles (angular dimensions of radius —30′) is considerably less than for each of them separately, if the distance between them does not exceed, for peripheric vision, (15° from the center) 95′, and for central vision, 5′. The main results obtained by Sevrungueena and myself are summarized in Table 1. The threshold was determined for one, two, or five points.

TABLE 1

	Red			Green			Blue		
Observers	1 point	3 points	5 points	1 point	3 points	5 points	1 point	3 points	5 points
I	100	36	12	100	86	45	100	93	70
II	100	49	20	100	72	68	100	80	70
III	100	36	24	100	83	15	100	58	42

*Accepted for publication by V. Borovski of the Editorial Board and received in the Editorial Office, April 1, 1935.

Each of these points was observed at an angle of 4′35″; the distance between the points was 41′12″ in the experiments with three points, and 18′22″ in those with five. The observation was conducted with the peripheric part of the retina (19° from the center). The color of the points was different: red, green, and blue. The threshold values for one point are accepted for 100, and those for three and five points are expressed in percentages of these values.

On the other hand Heymans (7) as early as 1901 demonstrated that the stimulus threshold for a round spot, the diameter of which is seen at the angle of 2°18′, increases in the presence of another luminous spot (4°34′ in diameter) and this increase in the threshold value is the greater, the brighter the second spot. The dependence of the inhibitive action of the second spot on its brightness (J) is expressed, according to Heymans, by the following formula, designated usually as "the Heymans law of inhibition":

$$R_J = R_o + hJ,$$

in which R_o is the threshold for the first spot in the absence of any other, R_J—its threshold in the presence of a second spot, the brightness of which is J, and h—a constant, which Heymans names "the coefficient of inhibition." Varying the distance between the spots within the range between 1°9′ and 10°12′ he found that with increasing distance this "coefficient of inhibition" decreases. Later Spencer (11) obtained, in experiments conducted on 50 persons, results corroborating the Heymans law of inhibition. However, Geldard (4) and other authors justly argued against these results, saying that the appearance of a second luminous spot in the field of vision causes the pupil to contract, and this contraction may be the actual cause of the threshold value increase. There can be no doubt that the effect observed by Heymans and Spencer depended to a certain degree on pupil contraction. But it is hardly possible to explain it by the action of the pupil reflex alone. The following considerations speak against such an assumption.

1. The coefficient h increases with the decrease in the distance between the spots, though the last named factor cannot increase the pupil contraction.

2. Spencer and Cohen (12) have found a highly positive correlation between the interindividual (50 persons) and the intra-

individual (one observer during 50 days) variations of the values R_o and h: in the first case the coefficient of correlation was +0.907, in the second +0.966. This fact is difficult to explain on the assumption that the value of h depends entirely on the pupil reflex.

3. McDougall (8, p. 64) observed the disappearance of a weak stimulus at the appearance of another stronger one in the field of vision even when an artificial pupil (1.5 mm.) was used.

The two groups of facts, on one hand those established by Beitel, Sevrugueena and myself, and on the other those found by Heymans and Spencer, seem to us not to be mutually exclusive. It is quite possible that the appearance in the field of vision of additional stimuli of liminal brightness lowers the threshold for the main stimulus, while additional stimuli of a brightness exceeding the liminal one raise this threshold. The present experiments were undertaken with the aim of verifying this supposition.

Technique

The inside walls of a light-impervious box (*A, B, C,* of Figure 1) are white diffusely reflecting surfaces lighted by lamp *L*. The box is placed in a perfectly dark room. The wall of the box facing

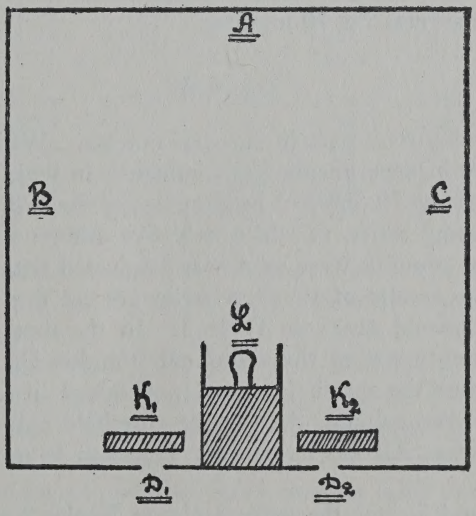


FIGURE 1

the observer has two round apertures D_1 and D_2. The observer sees them as two luminous circles the brightness of which may be varied by shifting the achromatic wedges K_1 and K_2.

The observations were conducted monocularly through an artificial pupil 4 mm. in diameter.

The threshold was always determined for the apertures D_1. The wedge K_1 was placed so that the observer would clearly see the spot corresponding to it. It was proposed to the observer to fix his eye on this spot. After this the brightness of the spot was gradually decreased by shifting the wedge, and the point of the wedge at which the spot disappeared was registered.

The threshold for D_1 was determined in the absence of additional stimuli (with the aperture D_2 closed) and in the presence of the additional stimulus D_2 the brightness of which was varied by corresponding shifting of the wedge K_2.

The angular distance between D_1 and D_2 was 65′, the diameter of the spots D_1 and D_2 in the first series of experiments (Tables 2 and 4) was 1′ 9″, in the second one (Tables 3 and 5), 2′ 18″. In all experiments D_1 was fixed with the eye, while the image of D_2 was outside the limits of the fovea but within those of the macula.

Before the beginning of each experiment the observers were kept in complete darkness for 70 minutes.

Results

Three persons took part in the experiments. With one of them the first series of experiments was conducted, in which the threshold was determined at 10 different brightnesses of the additional stimulus D_2. The second series, in which only five different brightnesses of the additional stimulus were used, was conducted with the other two persons. The results of the first series are set down in Table 2, those of the second series in Table 3. In the first column of the tables the brightnesses of the additional stimulus D_2 are registered, in the next five the thresholds for D_1, obtained in separate experiments, in the last column, the average threshold values for the person in question. All the figures are expressed in millionths of the stilbe.

Tables 4 and 5 show the same results as Tables 2 and 3, but here

TABLE 2

Brightness of additional stimulus	Thresholds found from single experiments No. 1	No. 2	No. 3	No. 4	No. 5	Average threshold value
			Observer I			
0.0000	0.5295	0.4284	0.3332	0.8151	0.6902	0.5593
0.1606	0.3808	0.4641	0.3332	0.7437	0.7021	0.5248
0.3213	0.3689	0.2975	0.2975	0.6723	0.9817	0.5236
0.6426	0.5593	0.2499	0.2320	0.7199	0.8508	0.5224
1.2733	0.3927	0.4343	0.2915	0.3927	0.7794	0.4581
2.6834	0.5057	0.3213	0.2439	0.4938	0.3927	0.3915
5.0337	0.5474	0.4105	0.3332	0.3451	0.5890	0.4450
9.9960	0.8568	0.5950	0.8925	0.5236	1.1483	0.8032
19.8551	0.6783	0.6664	1.1305	0.6485	0.6842	0.7616
28.0602	1.8504	1.6184	1.4458	0.7616	1.1721	1.3697
38.5500	2.7013	1.7612	1.8623	1.7969	1.4220	1.9087

TABLE 3

Brightness of additional stimulus	Thresholds found from single experiments No. 1	No. 2	No. 3	No. 4	No. 5	Average threshold value
			Observer II			
0.0000	0.4314	0.6063	0.6021	0.5832	0.3153	0.5076
0.1606	0.5176	0.2874	0.5688	0.4343	0.1904	0.3997
0.6426	0.3499	0.3034	0.3195	0.3748	0.2380	0.3171
2.6834	0.2660	0.2588	0.4153	0.3213	0.2201	0.2963
9.9960	0.4272	0.9109	0.5688	0.9579	0.5712	0.6872
38.5500	0.9591	1.2477	1.3858	1.6243	1.3149	1.3064
			Observer III			
0.0000	0.6081	0.7009	0.7009	0.8211	0.6366	0.6935
0.1606	0.3386	0.4373	0.7009	0.7021	0.6009	0.5560
0.6426	0.6468	0.5688	0.2874	0.5414	0.5771	0.5243
2.6834	0.4242	0.3195	0.5397	0.4343	0.3748	0.4185
9.9960	0.7414	0.8205	0.5397	0.2495	0.5410	0.9784
38.5500	1.5803	1.5387	1.1240	1.992	2.8857	1.8256

the threshold values obtained in the absence of the additional stimulus are designated as 100 and all the others are expressed in percentages of them.

The tables show that if the additional stimulus is relatively not very bright the threshold is lowered, while brighter additional stimuli raise the threshold. It follows that we may speak of two opposite effects of the additional stimulus: increase of the effect of the main stimulus (lowering of the threshold) and its inhibition (raising of the threshold), or in other words, of its positive and negative effect.

TABLE 4

Brightness of additional stimulus	Thresholds found from single experiments No. 1	No. 2	No. 3	No. 4	No. 5	Average threshold value
			OBSERVER I			
0.0000	100	100	100	100	100	100
0.1606	72	108	100	91	102	95
0.3213	70	69	89	82	142	90
0.6426	106	58	70	88	123	89
1.2733	74	101	87	48	113	85
2.6834	96	75	73	61	57	72
5.0337	103	96	100	42	85	85
9.9960	162	139	268	64	166	160
19,8551	128	156	339	80	99	160
28.0602	349	378	434	93	170	285
38.5500	510	411	559	220	206	381

TABLE 5

Brightness of additional stimulus	Thresholds found from single experiments No. 1	No. 2	No. 3	No. 4	No. 5	Average threshold value
			OBSERVER II			
0.0000	100	100	100	100	100	100
0.1606	120	47	95	74	60	79
0.6426	81	50	53	64	75	65
2.6834	62	43	69	55	65	59
9.9960	99	150	95	164	181	138
38.5500	220	206	230	278	417	270
			OBSERVER III			
0.0000	100	100	100	100	100	100
0.1606	56	62	100	85	94	79
0.6426	106	81	41	66	91	77
2.6834	70	45	77	53	59	61
9.9960	122	117	77	152	242	142

Under the conditions of our experiments a certain positive effect of the additional stimulus begins when its brightness is lower than the threshold, i.e., when the additional stimulus itself is not yet perceptible. With the increase in its brightness its positive effect is also increased until it reaches a maximal point—at the same brightness of 2.6834 for all the three persons tested, i.e., when the brightness of the additional stimulus is 4 to 5 times that of the threshold. Further increase in this brightness results in a decrease of the positive effect; at certain brightness values it becomes equal to zero, i.e., the additional stimulus does not exert any action on the effect produced

by the main stimulus. In our curves this "neutral point" corresponds to brightnesses varying for different observers between 6.300 and 6.600; in other words the additional stimulus did not exert any action on the effect of the main stimulus when its brightness exceeded the threshold stimulus 10 to 13 times. At still greater brightnesses the additional stimulus exerts a negative action, which increases with the increase of brightness. Our data do not suffice to answer the question as to what is the mathematical relation between the brightness of the additional stimulus and the value of its negative effect, nor in particular to make any conclusions in regard to the correctness or otherwise of Heymans' law of inhibition.

The experiments described bear a preliminary character. A preciser quantitative expression of the regularity found, as well as the elucidation of its dependence on a number of most important conditions—the distance between the stimuli, their location in relation to the point of fixation, and the time of exposition—constitute the subject of further experiments under way at the present moment. But already the above seems to us sufficient for affirming that the additional stimulus may, according to its brightness, exert both a positive and a negative action on the effect of the main stimulus. The similarity of the results obtained from experiments on different persons speaks for the reliability of the conclusions drawn.

The rule established by our experiments unites both groups of facts described in the beginning of this article, which seemed earlier to contradict one another, and, after a more detailed investigation, may throw light on the nature of the processes underlying the phenomena of the interaction of simultaneous light sensations.

Geldard in one of his recent works (5) made an attempt to prove that there is no inhibition effect within the limits of the fovea. However, the brightness of the additional stimulus in his experiments was only five times that of the main one. It is quite possible that with his technique as well a further brightness increase would result in converting the positive effect of the additional stimulus into a negative one. The fact that his curves show a maximum of the positive effect at a certain definite brightness of the additional stimulus and a decrease of this effect with further brightness increase speaks for the correctness of such a supposition.

In this connection it should be relevant to recall the experiments

of Graham and Granit (6), who showed that the additional stimulus does not increase the critical flicker frequency of the main stimulus if its brightness exceeds the brightness of the latter.

Summary

1. The action produced on the threshold of visibility of a luminous circle, fixed by the eye through an artificial pupil, by the presence (at a distance of 65′) of another luminous circle (additional stimulus) of varying brightness, has been studied.

2. The results showed that an additional stimulus whose brightness exceeds the threshold not more than 10 to 13 times exerts a positive action on the effect of the main stimulus (lowers its threshold), while a brighter additional stimulus acts negatively (raises the threshold) on the effect of the main one.

References

1. Beitel, R. J. Spatial summation of subliminal stimuli in the retina of the human eye. *J. Gen. Psychol.*, 1934, **10**, 311-327.
2. Charpentier, A. La perception des couleurs et la perception des formes. *C. r. Acad. Sci.*, 1883, **96**.
3. Fick, A. E. Studien uber Lichtsinn und Farbensinn. *Pflüg. Arch.*, 1888, **43**, 441-501.
4. Geldard, F. A. Brightness contrast and Heymans' law. *J. Gen. Psychol.*, 1931, **5**, 191-206.
5. ———. Flicker relations within the fovea. *J. Opt. Soc. Amer.*, 1934, **24**, 299-302.
6. Graham, C. H., & Granit, R. Inhibition, summation and synchronization of impulses in the retina. *Amer. J. Physiol.*, 1931, **98**, 664-673.
7. Heymans, G. Untersuchungen über psychische Hemmung. II. Artikel. *Zsch. f. Psychol.*, 1901, **26**, 305-382.
8. McDougall, W. Some new observations in support of Thomas Young's theory of light- and color-vision. *Mind, N. S.*, 1901, **10**, 52-97, 210-245, 347-382.
9. Piper, H. Uber die Abhängigkeit des Reizwertes leuchtender Objekte von ihrer Flächen—bzw. Winkelgrosse. *Zsch. f. Psychol.*, 1903, **32**, 98-112.
10. Sevruguenna, M. A., & Teplov, B. M. Dependence of the threshold value upon the number of incontiguous liminal stimuli. In *Visual Sensations and Perceptions.* Moscow, 1935 (Russian).
11. Spencer, L. T. The validity of Heymans' law. *Amer. J. Psychol.*, 1925, **36**, 427-433.

12. Spencer, L. F., & Cohen, L. H. The concept of the threshold and Heymans' law of inhibition. *J. Exper. Psychol.*, 1928, **11**, 88-97, 104-201, 281-292.

State Institute for Experimental Psychology
Moscow, U. S. S. R.

LA DÉPENDANCE DU SEUIL LUMINEUX DE LA PRÉSENCE DANS LE CHAMP VISUEL D'UN STIMULUS SUPPLÉMENTAIRE

(Résumé)

On décrit une expérience où l'on a étudié l'effet produit par la présence dans le champ visuel d'un stimulus supplémentaire sur le seuil absolu du stimulus principal. On avait pris pour stimuli principaux et supplémentaires des cercles lumineux dont les diamètres correspondaient à 1′ 9″ dans la première série, à 2′ 18″ dans la seconde. Le stimulus principal servait comme point de fixation; le stimulus supplémentaire se trouvait à une distance correspondante a 65′. On se servait d'une pupille artificielle. L'intensité lumineuse du stimulus supplémentaire était variable. On a trouvé qu'un stimulus supplémentaire, dont l'intensité lumineuse n'excède pas le seuil plus de 10-13 fois, produit un effet positif sur le stimulus principal (diminue le seuil), tandis qu'un stimulus supplémentaire d'une intensité lumineuse plus grande montre une influence négative (élève le seuil du stimulus principal).

Teplov

DIE ABHÄNGIGKEIT DES LICHTSCHWELLENWERTES VON DER ANWESENHEIT EINES NEBENREIZES IM SEHFELDE

(Referat)

Es wird ein Versuch beschrieben, worin der Einfluss, welcher die Anwesenheit eines Nebenreizes im Sehfelde auf die absolute Lichtschwelle für den Hauptreiz ausübt, studiert wurde. Als Haupt- und Nebenreize dienten Lichtkreise von 1′ 9″ (I Versuchsreihe) und 2′ 18″ (II Versuchsreihe) Durchmesser der Hauptreiz wurde durch eine künstliche Pupille fixiert; das Nebenreizbild lag peripherisch in einem Abstand von 65. Die Helligkeit des Nebenreizes wurde variert. Es ergab sich, dass ein Nebenreiz, dessen Helligkeit nicht mehr als 10-13 mal grösser ist als die des Schwellenwertes, einen positiven Einfluss auf den Hauptreiz ausübt (die Schwelle vermindert), während ein hellerer Nebenreiz eine negative Wirkung zeigt (die Schwelle für den Hauptreiz vergrössert).

Teplov

RECOVERY OF SENSITIVITY TO PRICK AND TOUCH AFTER PRESSURE BLOCK, WITH A NOTE ON THE SENSITIVITY OF HAIR SPOTS*

From the Psychological Laboratories of Harvard University

GERTRUDE RAFFEL

When pressure or anaesthetic block is applied to an excised nerve Gasser and Erlanger (2) have shown, from a study of action potentials, that the time of loss of function is related to the size of nerve fibers: the large fibers are affected before the small ones by compression, and the smaller fibers are the first to be blocked by cocaine. When the block is removed, the order of recovery is the reverse of the order of loss of conductivity. A series of experiments by Heinbecker, Bishop, and O'Leary (5, 6, 7) is concerned with the sensitivity of cutaneous nerves. They relate fiber size to the phenomena of sensation, and report that the effects of novocaine block occur in the following order: loss of sensitivity to heat and cold, then to cutaneous (prick) pain and to pressure pain, loss of the motor function, of the joint sense, of sensations of pressure, and finally of sensations of touch. The action currents in these studies show potential waves with maxima along a continuous distribution; the authors believe, accordingly, that fibers occur in groups according to their terminal connections, that these groups vary in size and conduction rate, but that there is overlap between them. The immediate implication of the overlap is that, when, for example, thermal sensitivity is lost through novocaine block, sensitivity to pain has also been diminished, and there is probably some impairment of motor function. The total picture of sensitivity, at any period during application of a block or during recovery from it, is not, therefore, a simple one.

This experiment traces recovery of sensitivity to prick and to touch after removal of compression block in the intact organism.

*Recommended by Edwin G. Boring, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, April 1, 1935.

The method of applying the block is essentially the one that was used in early experiments by Herzen (8) and Goldscheider (3), who studied the order of disappearance of warmth, cold, touch, and pain sensitivity when a leg or an arm had "fallen asleep." The experiment differs from previous ones in that it determines the course of the changes in sensitivity during recovery, rather than reporting only the time at which recovery is complete.[1]

Procedure

The area used for experimentation was on the dorsal surface of the left leg, 16 cm. above the heel. Control series were taken from the same area, or from the corresponding region of the right leg. The writer acted as subject.

The block was applied by crossing the left leg over the right so that the right knee pressed against the superficial peroneal and sural nerves, where they run parallel across the popliteal fossa. The first hint of numbness appeared in the toes after about 11 minutes; after 20 to 30 minutes it had spread to the experimental area. When the area was insensitive to a cold object passed lightly over it the block was removed. Readings were begun about 10 seconds after removal of the block, and continued for 30 minutes.

The usual complex of tingling, ache, and waves of heat and cold, which would mask any stimuli except extremely intense ones, are reduced if the leg is not moved after the period of block, and are virtually absent if it is not moved and is held horizontal. During the period of recovery the subject lay with the dorsal surface of the leg upward; distracting sensations were so sporadic and transitory as hardly to act as distractions. In a typical session, for example, there was one wave of cold in the second and one in the fifth minute after the block was removed; tingling was felt in about the fortieth second; there was some pain during the second and third minutes; and a feeling as of relaxing muscles in the second minute and the fifth. Each of these lasted for less than a second.

A limen gauge algesimeter similar to the one described by Boring (1) was used as stimulus for prick. It recorded pressures from

[1]The writer is glad to acknowledge her gratitude to Dr. Edwin G. Boring for his advice throughout the experiment.

zero to seven grams. An area on the leg 2 cm. square was stamped in millimeter squares; successive stimulations were given in adjacent squares along the longitudinal axis of the leg. Preliminary experiments demonstrated that there was no difference in sensitivity between the hair spots and the surrounding areas.

The procedure was as follows. The experimenter said "Ready," touched the point of the algesimeter to the skin, and gradually increased pressure upon it. As soon as pain was felt the subject said "Now," the experimenter gave the reading of the gauge, the subject recorded it, the experimenter said "Ready," and began the next stimulation. A clock faced the subject, who recorded on the data sheet when each minute had passed. Since pressure was applied for a longer time when the threshold was high, the number of stimuli given in each minute was variable. There were never less than seven stimuli in a minute; the average number was 10. The procedure in the control series was precisely the same, except that no block had previously been applied. Several hours of preliminary work familiarized both the experimenter and the subject with the procedure.

Since the limen gauge did not measure pressures greater than seven grams, and since occasionally a pressure of seven grams did not cause pain, certain of the distributions of thresholds are unclosed. The median was therefore used as the measure of central tendency. The data are presented graphically; each point on the curve represents the median of the limena obtained during a two-minute period. The data of two experimental sessions are averaged in Figure 1; in Figure 2 the median thresholds of a single control session are shown, and also the average of this series and of another control session.

A Von Frey aesthesiometer with a moderately thin horse hair was used as stimulus for touch sensitivity. The length of the hair could be varied between 0 and 4 cm.; the intensity of the stimulus was inversely proportional to the length of the hair. The area to be used was shaved before each series; experimentation was never begun until after the tingling from the shaving had stopped. The area was stamped as before. Care was taken not to stimulate a hair spot; if a hair spot was touched accidentally, the reading was discarded. Stimuli were applied in adjacent millimeter squares in time with a

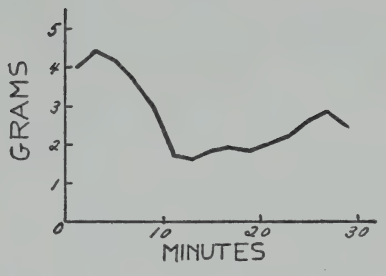


FIGURE 1

RECOVERY OF PRICK SENSITIVITY AFTER REMOVAL OF PRESSURE BLOCK

Showing changes in threshold for 30 minutes after removal of the block.

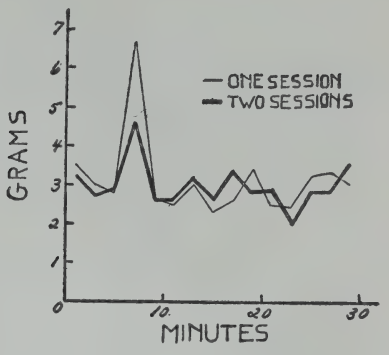


FIGURE 2

CHANGES IN PRICK SENSITIVITY FOR 30 MINUTES UNDER NORMAL CONDITIONS

metronome which rang 14 times a minute; the bell of the metronome acted as a ready signal. The subject recorded whether or not the stimulus was felt. As with work on prick, there were several hours of preliminary experimentation.

The data are presented in terms of the threshold for successive periods of two minutes, obtained from two experimental sessions. The control series represents only one session; it may therefore be compared with the curve of the single control series for prick sensitivity.

During one experimental series the hair length was adjusted so that pressure was judged present more than 50 per cent of the time; in the other experimental series it was judged present in less than 50 per cent of the cases. Thresholds were obtained from these values. The thresholds of the control series were found by the following method:[2] a threshold stimulus, previously determined, was used throughout. The percentage of judgments of present in each two-minute period was calculated. On the assumption that the psychometric function is an integral of the normal probability curve, each of these percentages was transformed into a standard score in terms of its deviation from 50 per cent. The 50

[2]Suggested by Dr. John Volkmann.

per cent point of the psychometric function for which the stimulus length would have this value was then obtained.

Results

Three series were taken to determine the course of recovery of prick sensitivity after nerve block; one of these sessions lasted for 12 minutes; the other two lasted for 30 minutes. The data of the latter two were averaged and are presented in Figure 1. All three curves showed the same characteristics: a high initial threshold, a slight rise, which reached its peak in three or four minutes, and a rapid drop to a value far below the normal threshold. In the two series in which experimentation was continued, the curves rose gradually, until at about 25 minutes they showed random fluctuations around the normal value. Although the decrease in sensitivity during the first few minutes was slight, the fact that it was present in every session indicates that it is significant.

It is clear that under normal conditions only random irregularities are present in prick sensitivity, as shown in Figure 2. The median threshold for the first half of the period is 2.8 gm., and for the second half, 2.6 gm. The fluctuations are relatively large, and apparently are not rhythmical.

The control series for touch corresponds closely to that for prick; there are relatively large non-rhythmic fluctuations in both Figure 2 and Figure 4. The mean threshold stimulus for the first 15 minutes of the period was 1.64 cm.; for the second 15 minutes it

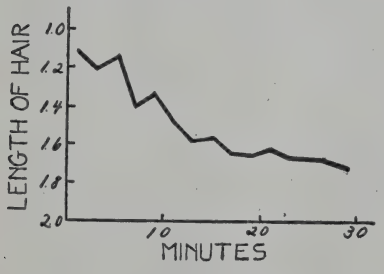


FIGURE 3
Recovery of Touch Sensitivity after Removal of Pressure Block
Showing changes in threshold for 30 minutes after removal of the block.

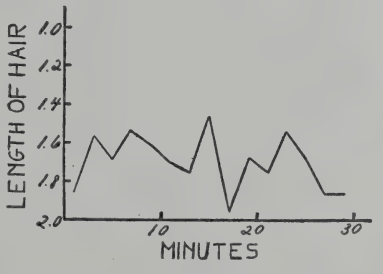


FIGURE 4
Changes in Touch Sensitivity for 30 Minutes under Normal Conditions

was 1.76 cm. In Figure 3 is presented the curve for recovery of touch sensitivity after blocking. It shows a high initial threshold, and a gradual increase in sensitivity, broken by minor fluctuations, until at about 25 minutes after removal of the block the threshold is normal. The curve is negatively accelerated except in its initial portion, which is flattened.

Discussion

The curve of recovery of prick sensitivity after blocking shows two unexpected characteristics: the rise in threshold during the first few minutes, and the period of hyperalgesia between 10 and 25 minutes. Since neither of these characteristics is present in the curve of recovery of touch sensitivity, they cannot be regarded as typical changes in peripheral fibers after nerve block. The writer suggests that both the initial and the secondary rise in the thresholds of prick are due to masking from other groups of fibers.

According to Heinbecker, Bishop, and O'Leary, temperature fibers should be the first to recover from compression block, then pain fibers, and, after a longer interval, the touch fibers. The data obtained in the present experiment for pain and touch are consistent with such an account: thresholds for prick fall at or below the normal value after nine minutes; and thresholds for touch have reached normal at about 25 minutes after removal of the block. The first increase in the prick threshold reaches its maximum in three minutes, the time at which recovery of the temperature fibers should be far advanced. The period of hyperalgesia comes to an end at approximately the time that recovery of the touch fibers is complete. Such relationships suggest that thermal and touch sensitivity mask sensitivity to prick.

In line with this hypothesis, the following description of the recovery of prick sensitivity after pressure block may be given. During the first three minutes after removal of the block, temperature fibers recover more rapidly than pain fibers; the masking from temperature decreases sensitivity to prick more than the recovery of pain fibers increases it, and the curve of Figure 1 rises. After three minutes most of the temperature fibers have recovered, and there is little increase in masking; thus increased sensitivity to prick due to the recovery of pain fibers becomes apparent, and the curve

falls. This is the period of most rapid recovery of pain fibers; probably almost all of them have recovered after 13 minutes. Touch fibers continue to recover after this time, and thus their masking of pain increases. Since there is no further gain in sensitivity for prick to offset the masking, the threshold for prick rises for the second time.

Support for the hypothesis of masking may be found in the normal variability of the thresholds for touch and prick. Such large changes as are shown in Figures 2 and 4 might be expected to occur if there were interaction between different afferent systems, resulting in an unstable balance in the spinal centers. If the theory is valid, it shows that, as was suggested by Head (4), it is possible for touch to inhibit pain.

The Sensitivity of the Hair Spots

During preliminary experimentation it was found that hair spots are not more sensitive to prick than are other areas, but that, even when the hairs are shaved, they are more sensitive to touch (Table 1). Stetson (9) has removed the hair follicle, leaving the hair in place, and found almost no change in threshold; the limen for pressure was slightly lowered after the operation. He concluded that the superior sensitivity of the hair spots is due to the fact that the hair when stimulated acts as a lever, deforming the skin. According to this hypothesis, if the hair is removed completely, there should be no greater sensitivity at the hair spot than elsewhere on the skin.

In a supplementary experiment, the hairs of a small area on the subject's leg were removed with tweezers. The area was examined under a magnifying glass to make certain that no hairs

TABLE 1

	Limen
Prick	
Hair spots	2.7 gm.
Other areas	2.8 gm.
Touch	
Hair spots, hairs shaved	1.35 mm.
Hair spots, hairs removed	1.3 mm.
Other areas	0.7 mm.

remained. The hair spots were stimulated under these conditions (Table 1). The aesthesiometer was equipped with a human hair for the judgments; therefore the liminal values cannot be compared to those previously obtained.

Removal of the hairs does not reduce the sensitivity of the hair spots to that of the rest of the skin. Stetson's explanation of the sensitivity of the hair spots thus appears to be inadequate.

References

1. Boring, E. G. Cutaneous sensation after nerve-division. *Quart. J. Exper. Physiol.,* 1916, **10**, 1-98.
2. Gasser, H., & Erlanger, J. The rôle of fiber size in the establishment of a nerve block by pressure or cocaine. *Amer. J. Physiol.,* 1929, **88**, 581-591.
3. Goldscheider, A. Zur Dualität des Temperatursinns. *Pflüg. Arch.,* 1886, **39**, 96-120.
4. Head, H., & Rivers, W. H. R. A human experiment in nerve-division. *Brain,* 1908, **31**, 323-450.
5. Heinbecker, P., Bishop, G. H., & O'Leary, J. Allocation of function to specific fiber types in peripheral nerves. *Proc. Soc. Exper. Biol. & Med.,* 1932, **30**, 394-405.
6. ————. Analysis of sensation in terms of the nerve impulse. *Arch. Neur. & Psychiat.,* 1934, **31**, 34-54.
7. ————. Pain and touch fibers in peripheral nerves. *Arch. Neur. & Psychiat.,* 1933, **29**, 771-789.
8. Herzen, A. Über die Spaltung des Temperatursinnes in Zwei gesonderte Sinne. *Pflüg. Arch.,* 1885, **38**, 93-103.
9. Stetson, R. H. The hair follicle and the sense of pressure. *Psychol. Monog.,* 1923, **32**, 1-17.

Monmouth Junior College
Long Branch, New Jersey

LA RESTITUTION DE LA SENSIBILITÉ À LA PIQÛRE ET AU TACT APRÈS LE BLOCAGE PAR PRESSION, AVEC UNE NOTE SUR LA SENSIBILITÉ DES POINTS OÙ SE TROUVENT UN CHEVEU

(Résumé)

On a appliqué le blocage par pression aux nerfs suraux et péroniers superficiels jusqu'à ce que la jambe "s'est endormie," et que l'on a observé la restitution de la sensibilité à la piqûre et au tact. La sensibilité au tact montre une diminution graduelle du seuil; la restitution est complète après 25 minutes. La sensibilité à la piqûre décroît pendant 3 minutes

après l'enlèvement du blocage; ensuite elle augmente rapidement pendant 8 minutes de sorte que la peau est hyperalgésique; et ensuite elle décroît graduellement jusqu'à ce que le seuil normal est approximé à 25 minutes. L'auteur suggère que les deux périodes de sensibilité augmentante sont dues au recouvrement de la piqûre comme les fibres de température et de tact se récupèrent.

On a constaté que, même après l'enlèvement complet du cheveu, le point où se trouve le cheveu est plus sensible au tact que la peau environnante.

RAFFEL

DIE WIEDERHERSTELLUNG DER EMPFINDLICHKEIT AUF STECHEN UND BERÜHRUNG NACH EINER DRUCKABSPERRUNG, MIT EINER ANMERKUNG ÜBER DIE EMPFINDLICHKEIT DER HAARFLECKE

(Referat)

Die Druckabsperrung wurde auf die Waden- und oberflächlichen Wadenbeinnerven aufgelegt, bis das Bein "einschlief," und die Wiederherstellung der Empfindlichkeit auf Stechen und Berührung wurde beobachtet. Die Berührungsempfindlichkeit weist eine allmähliche Vermehrung der Schwelle auf; die Wiederherstellung wird nach 25 Minuten vollendet. Die Stechenempfindlichkeit nimmt für 3 Minuten nach dem Wegschaffen der Sperrung ab; dann nimmt für 8 Minuten schnell zu, so dass die Haut überempfindlich ist; und dann nimmt allmählich ab, bis die normale Schwelle nach 25 Minuten beinahe erreicht wird. Der Autor schlägt vor, dass die zwei Perioden der abnehmenden Empfindlichkeit die Folge der Verdeckung des Stechens sei, wie die Temperature- und Tastfasern sich weiderherstellen.

Es fand sich, dass der Haarfleck empfindlicher auf Berührung ist als die angrenzende Haut, sogar nach dem vollkommenen Wegschaffen der Haare.

RAFFEL

LURIA'S MOTOR METHOD AND WORD ASSOCIATION IN THE STUDY OF DECEPTION*[1]

From the Department of Psychology, University of Wisconsin

JOSEPH E. RUNKEL

I. INTRODUCTION

The controlled word-association method as established by Galton and improved upon by Wundt, Jung, and others has been frequently applied to the study of deception. The results obtained by Yerkes and Berry (15) and Wertheimer and Klein (16) seemed to indicate positive results. Crane (7) in 1915 declared that the word-association method in itself was not a reliable indicator. Crosland (8) criticized Crane's method, and conducted an experiment where the method was successful with the use of extensive statistical application. Two of the more important criticisms that have been leveled against this method are (1) that the sophisticated subject may prearrange his response so as to make his reaction time uniformly long to all of the words, and (2) minute variations in the reaction time are often insignificant in indicating deception, so that it is difficult to determine how great a difference is necessary to be considered significant.

Physiological measurements have been made in combination with the word-association technique in many experiments. The first attempt to use the pneumograph as an indicator of physiological changes accompanying emotional arousal in deception was made by Benussi (1), who used the direct question method. He found that the ratio of inspiration to expiration or I/E was less before lying than after lying.

Burtt (4) also used this method, and found that the positive

*Recommended by Hulsey Cason, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, April 2, 1935.

[1]The writer wishes to acknowledge the assistance and encouragement of Professor Hulsey Cason in carrying out the present study.

results were 20 per cent better than chance. Marston (12) recorded changes in blood pressure in deceptive situations and on the basis of rises of blood pressure judged whether the subject was deceiving or not. Larson (11) adopted the fundamental assumptions of Marston and refined them so that the blood pressure was recorded continuously. Larson's method has been used in practical application, and successful results have been obtained. Burtt (4) states that "The successful application of psychology to the problems of deception will involve a combination of several methods. Suspects examined by such methods would then fall into three groups—those in which all the measures point towards guilt, those in which all measures point towards innocence, and those in which the measures disagree."

The study of deception is inextricably connected with the study of affective processes, and any objective indicator that is used will in reality indicate in some manner the behavior changes that take place in the organism. For the most part, the research in the field of affective processes has been concerned with the recording of peripheral changes caused by and involved in the affective processes. Typical studies have been concerned with the recording of blood pressure by Marston (12), Landis (10), and Blatz (2); respiratory changes by Benussi (1) and Burtt (4); gastro-intestinal changes by Brunswick (3) and others; blood pH changes by Starr (14); salivary pH by Gaskill (8); and output of adrenalin by Cannon (5) and his colleagues.

It is the contention of Luria (13) that emphasis should be shifted from the peripheral processes to the central processes and that studies should be made of the disorganization of behavior. He says:

> In thus shifting our emphasis to the central processes we are in step with the leading physiologists. The researches of Cannon . . . showed experimentally that the complete excision of the viscera did not result in a decrease of emotion, and that the animal after the operation gave evidence of all the signs of affective behavior in spite of the fact that the visceral components were absent. These experiments confirm that the physiological changes, usually considered as basic for affective processes, in reality are only secondary and accompanying processes.

According to Luria, the arousal of an affect basically changes the structure of the reactive processes. The affect modifies the higher cortical restraint, the voluntary control of a motor response which is mediated through these higher centers is lessened, and fine motor adjustments are rendered more difficult. The affect does more than exert an inhibitory influence on the cortex; it causes a diffusion of the impulses and changes the normal reactive pattern. Luria believes that it is desirable to use an active voluntary motor response which is connected with the language mechanism. The change that the affect produces in the central nervous system may be recorded in the change of the reactive structure. He used the word association method and motor response simultaneously. The inhibitory powers of the cortex were brought into play in the establishment of a mental set between the language stimulus and the motor response. The subject had to inhibit a response until he replied with the first word that came to mind. When the *St.* (stimulus) word aroused an affective tone sometimes a lessening of cortical restraint occurred, which tended to break down the inhibition established, and at the same time to produce a slight motor tremor.

The experiment herein described was designed to create as satisfactory a laboratory deceptive situation as possible, and then to give the subjects a word-association test in conjunction with the motor method. We noted changes in motor reactions and the verbal reaction time. Burtt has made a study of "Motor Concomitants in Word Association," and has kindly let me read the manuscript of his paper, which has been published in the February, 1936 issue of the *Journal of Experimental Psychology.*

II. Materials and Apparatus

The apparatus consisted of two main sections: (1) a pneumatic system for measuring the motor response, and (2) apparatus used in measuring reaction time in the word-association test. The pneumatic system consisted of two oblong rubber bulbs 3 in. by 3½ in. x 1 in., which were constructed of thin rubber. These bulbs were held firmly in place by wooden frames whose internal dimensions were the same as the external dimensions of the bulbs. The rubber tubing which connected the two bulbs with two Marey

tambours mounted on a Stoelting universal support was approximately 50 in. long and had an internal diameter of 2/16 in. The recording arms on the tambours amplified the excursion six times and the tambours were mounted on the universal support so that one writing point was directly above the other. These two arms recorded on the smoked paper of an extension kymograph.

The device used for measuring the word-association reaction time consisted of a stop watch in a stop watch controller connected in circuit with a telegraph key and a signal marker. The signal marker was connected so that when contact was made with the telegraph key, the stop watch controller started the stop watch and simultaneously activated the signal marker. A second telegraph key in a separate electric circuit was used to reset the stop watch. A Jacquet graphic chronometer recorded the time in seconds on the smoked paper which moved at a speed of approximately 20 cm. per second.

A wooden screen composed of a center panel 22 in. by 25 in. and two oblique wing panels 16½ in. by 15 in. were placed in front of the subject in order to conceal the apparatus. In the center panel there were two apertures 5 in. and 7 in. apart, through which the *S* (subject) placed his hands upon the bulbs.

III. Preliminary Series of Experiments

The procedure consisted of two parts: a preliminary experimental series in which five subjects were used, and a main experimental series in which eight subjects were used. The preliminary experimental series was divided into three main sections: The Pre-Test, the L.D. Situation (laboratory deceptive situation), and the Final test. In the word-association procedure in the Pre-Test each *S* was given 100 words, one at a time. Eighty of these 100 words were taken from the Kent-Rosanoff list, and these 80 words were our N-C (non-critical) words. The 20 words that had critical connotations after the deceptive situation were chosen from experiences the subject had during the laboratory deceptive situation.

Pretest. After *S* had been seated and his hands placed upon the bulbs, he was told that he would be given a word and that he must respond with the first word that came to mind, and at the same time must press lightly with his right hand on the right-hand bulb.

S was instructed to let his left hand rest passively on the left-hand bulb, and a record was made of any movements in the left hand because it was thought that a diffusion of nervous impulses might influence the left hand and cause a response. As *E* (experimenter) gave the stimulus word, he simultaneously pressed the telegraph key which started the stop watch. When *S* responded, *E* stopped the watch by means of the key. The motor record and the word-association record were used in the Pre-Test to determine whether or not any of the words were complex indicators.

Laboratory Deceptive Situation. After obtaining the Pre-Test record, *S* was mailed a sealed set of directions which read as follows:

> Go to the side door of Bascom Hall that is located on the north side of the building at 11 P.M. tonight. Enter the building and go towards the stairs which lie directly in front of you. About 5 feet from the bottom of the stairs you will note a narrow hallway leading to the left. Turn down this hallway and enter the first room to the left; there will be a curtain drawn over the entrance. On a ledge to the right you will note a black box. Open this box and remove an envelope containing further directions.

The following considerations were employed in attempting to devise a satisfactory laboratory deceptive situation: (1) The situation should arouse affective processes which will be later recalled by the *C* stimulus word. (2) It is desirable that the objects which aroused fear or startle reactions in *S* be clearly and vividly perceived. (3) It is desired to make these impressions through as many senses as possible. With these general purposes in mind, the following situation was devised.

A stained glass window separated the room which *S* was to enter from an adjacent room where *E* observed *S*'s behavior. Because of the total darkness of the observation room in contrast with some light in *S*'s room, it was possible to observe *S* without his being aware of anyone in that portion of the building. In order for *S* to reach this room it was necessary for him to go through 20 feet of darkened hallway. The box referred to in the above directions contained an electric bulb connected with a switch in the observation room and *E* could flash on the light just as *S* opened the box. The instructions in the box were as follows:

> Remove the sheet on the ledge to the left. Take the knife that is placed by the animal and cut off a leg of the animal. After doing this, place the sheet over the animal and leave the building. Do not speak to anyone concerning these occurrences.

S cut off the leg of a dead rat that had been anaesthetized a short time before. Just as *S* was cutting off the rat's leg, a bottle placed on a ledge directly behind the *S* was pushed off by means of an electro-magnetic device. The bottle made a loud sound as it broke against a piece of iron on the floor. Just as *S* started to leave the room, a shot was fired and a girl screamed. Everything was then quiet until *S* left the building.

Final Test. The final test occurred 24 to 36 hours after the Laboratory Deceptive Situation. *S* was seated and was told that *E* was going to try to find out what he had done the previous evening. *S* was asked to try to conceal any clues from *E* that he could. The Pre-Test procedure was then repeated. The 20 words, that had no critical connotation before the deceptive situation, now had critical implications. The main purpose of the preliminary series was to discover the effectiveness of the critical words and to determine the value of the N-C words as minimal complex indicators. The preliminary experiment served as a control for the main experiment.

Main Series of Experiments. The main series of experiments was designed to correct some of the faults found in the preliminary series. In the preliminary experiments, the Pre-Test gave *S* a clue as to what to expect in the Final Test. In the Final Test several *S*'s were able to remember the responses given in the Pre-Test, and they knew what the critical words would be. The Pre-Test was therefore abandoned in the main series of experiments. The long delay between the Laboratory Deceptive Situation and the Final Test was also shortened, and *S* was tested a few minutes after the Laboratory Deceptive Situation. In the main series of experiments the words in the association list were reduced from 100 to 35 and 10 of these were critical words. We selected the 10 C (critical) words that seemed of most value in the preliminary experiments. The 25 N-C words selected were in general the least indicative of complexes.

The instructions used in the Laboratory Deceptive Situation in the main series of experiments were as follows:

> Take the candle on the ledge to the right of the box. Light it and go to room 19, where you will see a box on the table. Open the box, and carefully examine the contents. Then leave the room through the rear window, and leave the grounds by the same way you came.

In the main series of experiments the Laboratory Deceptive Situation was changed as follows: A human skull and two pictures of decaptitated soldiers were substituted for the rat. *S* lit the candle and examined the skull and pictures in candle light. A bell rang just as *S* began examining the skull, and a shot was fired immediately after the bell.

When *S* left the building through the rear window, he was apprehended by an assistant who said nothing to him, but led him into the building and placed him in a lighted room to wait a short time before the Final Test. In this procedure *S* did not know what would happen next, and he could not fortify himself for the test situation. He was also unfamiliar with the apparatus and had no way of planning his responses.

IV. Results

Preliminary Experiments. In the preliminary experiments, results were obtained on five male university students. In calculating the results we compared the mean and median reaction times to the C and N-C words before and after the Laboratory Deceptive Situation. We examined the motor records of the hand movements in response to each of the stimulus words, and determined whether or not any definite tremors occurred in the records of each hand. The records of the left hand were not found to be indicative and they were not included in the calculation of the results. Using a response of the right hand to each stimulus word as one trial, the average number of trials in which a definite tremor occurred was 6 per cent in the Pre-Test and 28 per cent in the Final Test. The fact that tremors were present in only 6 per cent of the trials in the Pre-Test seems to indicate that only a few of our words had a definite value as complex indicators at the beginning of the experiment.

The mean and median results on reaction time in response to the stimulus words in the Pre-Test and Final Test are given in Table 1.

TABLE 1

REACTION TIME TO STIMULUS WORDS IN THE PRE-TEST AND FINAL TEST

		Mean in sec.		Median in sec.	
		20 C Words	80 N-C Words	20 C Words	80 N-C Words
Pre-Test	1	2.5	2.0	1.8	1.8
	2	1.8	1.9	1.8	2.0
	3	2.6	2.1	2.0	2.0
	4	2.0	1.9	2.0	1.8
	5	1.5	1.9	2.2	2.0
Final Test	1	2.6	2.0	2.2	1.8
	2	2.0	1.9	2.0	2.0
	3	1.7	1.8	1.8	1.8
	4	2.5	2.0	2.2	2.0
	5	2.4	2.0	2.4	2.0

The reaction time to the 20 C words and the 80 N-C words are in adjacent columns. In each group of five figures in Table 1, the first figure was obtained with the first *S*, the second with the second *S*, etc. In both Pre-Test and Final Test, the speed of response to the C and N-C words was practically the same. From the results obtained in the main series of experiments, it seems that the use of the Pre-Test gave the *S* some information in regard to the nature of the Final Test. Because of this information *S* was able to control his reactions in the Final Test. *S* No. 3 in Table 1 told *E* that after the Pre-Test he planned the responses he would give if tested again.

Table 2 gives the percentage of trials where tremors occurred in

TABLE 2

PERCENTAGE OF MOTOR TREMORS TO 80 NON-CRITICAL AND 20 CRITICAL WORDS AFTER THE LABORATORY DECEPTIVE SITUATION

Subjects		1	2	3	4	5
Percentage of motor tremors	To 80 N-C Words	11.0	2.2	18.7	32.5	37.5
	To 20 C Words	20.0	10.0	15.0	60.0	50.0

the right hand record to the C and N-C words after the Laboratory Deceptive Situation. In four *S*'s out of five, there was a higher percentage of tremors to the C words than to the N-C words. The purpose of this preliminary series was to standardize the procedure, and no detailed analysis of the results is attempted.

Main Experiment. The results of the main series of experiments were obtained with eight university students, seven of whom had had some experience with the word-association method in a course in experimental psychology. The 25 N-C words used in the present experiment were standardized from the 80 N-C words used in the preliminary experiment.

The average reaction time of the five *S*'s to the 25 N-C words in the Final Test in the preliminary series was 1.8 sec. with an A.D. of .28 sec. The 10 C words in the present experiment were chosen as far as possible on the basis of introspections of the *S*'s in the preliminary experiments. We also attempted to choose common words for which *S* had a ready association, so that the word in itself would not lengthen the reaction time. Words in common usage such as shot, gun, box, bell, and the like were chosen as far as possible.

TABLE 3

REACTION TIME TO 25 NON-CRITICAL AND 10 CRITICAL WORDS

Subject	Mean		Median	
	N-C	C	N-C	C
E.S.	1.9	3.9	2.0	3.8
J.M.	1.6	2.7	1.5	2.1
R.D.	1.5	2.1	1.5	2.1
R.S.	1.4	2.7	1.4	2.0
G.C.	1.4	2.5	1.4	2.0
E.K.	1.5	2.3	1.5	2.1
J.K.	1.2	1.6	1.3	1.7
E.H.	1.6	2.3	1.5	1.9

Table 3 shows the reaction time of the eight *S*'s to the 25 N-C and the 10 C words. The reaction time to the C words was longer than to the N-C words for both the mean and the median in the case of every subject.

Table 4 gives the detailed results on the reaction times and verbal responses of the eight *S*'s to each of the 10 C words. The row labeled "Av. R.T. 25 N-C Words" gives the average reaction time to the 25 N-C words for each *S*. Out of the total of 78 C words (two being omitted) there were 10 which had a shorter reaction time than the average of the N-C words. The words in italics are the responses which we regarded as truly indicative. The reac-

TABLE 4

REACTION TIMES AND VERBAL RESPONSES TO THE 10 CRITICAL WORDS

Subject / St. Word	E.S. R.T.	E.S. Response	J.M. R.T.	J.M. Response	R.D. R.T.	R.D. Response	R.S. R.T.	R.S. Response
Skull	2.0	Bones	1.3	Bones	1.8	Head	1.9	Head
Shot		(Omitted)	1.5	Gun	2.0	Hurt	1.9	Hit
Gun		(Omitted)	2.0	Rifle	2.1	Shot	2.0	Shoot
Deceive	*8.0*	*Control*	2.0	Fool	*2.7*	*Perceive*	*3.0*	*People*
Candle	*3.8*	*Lamp*	*3.7*	*Pump*	1.9	Light	1.8	Light
Bell	*3.8*	*Sound*	1.6	Ring	1.4	Ring	1.1	Ring
Flash	2.2	Light	2.1	Light	1.8	Light	1.2	Light
Box	*4.0*	*Can*	*4.1*	*Square*	2.1	Wood	*4.0*	*Wood*
Decapitate	*5.8*	*Captain*	*4.0*	*Behead*	*3.0*	*Head*	*8.0*	*Brick*
Bones	2.1	Skull	*5.1*	*Skeleton*	2.2	Foot	2.3	Break
Av. R.T. 25 N-C Words	1.9		1.6		1.5		1.4	

Subject / St. Word	G.C. R.T.	G.C. Response	E.K. R.T.	E.K. Response	J.K. R.T.	J.K. Response	E.H. R.T.	E.H. Response
Skull	1.6	Box	1.9	Black	*1.8*	*Head*	1.4	Box
Shot	2.0	Bank	1.5	Gun	1.6	Hard	1.5	Dark
Gun	1.3	Shot	2.2	Bullet	1.5	Man	1.5	Man
Deceive	*9.0*	*Man*	2.2	Conceive	*1.7*	*Man*	*5.1*	*Bought*
Candle	2.0	Box	1.8	Light	1.5	Light	1.9	Light
Bell	2.0	Gong	2.0	Ring	*1.7*	*Ring*	1.4	Ring
Flash	2.0	Light	*3.2*	*Light*	1.3	Light	1.9	Sound
Box	1.9	Skull	2.1	Candle	1.4	Hard	*2.5*	*Candle*
Decapitate	1.2	Behead	*4.5*	*Part*	*1.9*	*Head*	1.9	War
Bones	2.1	Head	*5.1*	*Hard*	*1.7*	*Head*	*3.8*	*Stones*
Av. R.T. 25 N-C Words	2.5		2.3		1.6		2.3	

tion times to these italicized words are greater than the average reaction time to all of the remaining C words given in the table.

Table 5 gives the percentage of truly indicative verbal responses

TABLE 5

PERCENTAGE OF TRULY INDICATIVE VERBAL RESPONSES AND INDICATIVE MOTOR RESPONSES TO THE 10 C WORDS

Subject	Percentage of truly indicative verbal responses	Percentage of indicative motor responses	Average percentage
E.S.	50	80	65
J.M	40	30	35
R.D.	20	70	45
R.S.	30	60	45
G.C.	10	60	35
E.K.	30	40	35
J.K.	50	20	35
E.H.	30	40	35
Av. percentage	32.5	50	41.2

and indicative motor responses to the 10 C words for each of the eight *S*'s. In the case of two of the eight *S*'s the percentage of truly indicative verbal responses was greater than the percentage of indicative motor responses.

The eight *S*'s gave a total of 26 truly indicative verbal responses and a total of 40 indicative motor responses out of a possible 160 responses to the C words. For all *S*'s, 32.5 per cent of the verbal responses to the C words were indicative, and 50 per cent of the motor responses to the C words were indicative. The last column in Table 5 shows that, out of a possible 20 verbal and motor responses, subject E.S. gave 65 per cent indicative responses while subjects E.H., J.M., G.C., and J.K. each gave 35 per cent indicative responses.

An examination of the combined results of all the *S*'s not given in Table 5 shows that words which tend to call out a truly indicative verbal response also tend to call out an indicative motor response. A rank correlation of .40 was found between the frequency of truly indicative verbal responses and the frequency of indicative motor responses to the 10 C words.

In Figure 1 are shown representative sections of the motor records of six of the *S*'s. The time line, the verbal stimulus and verbal

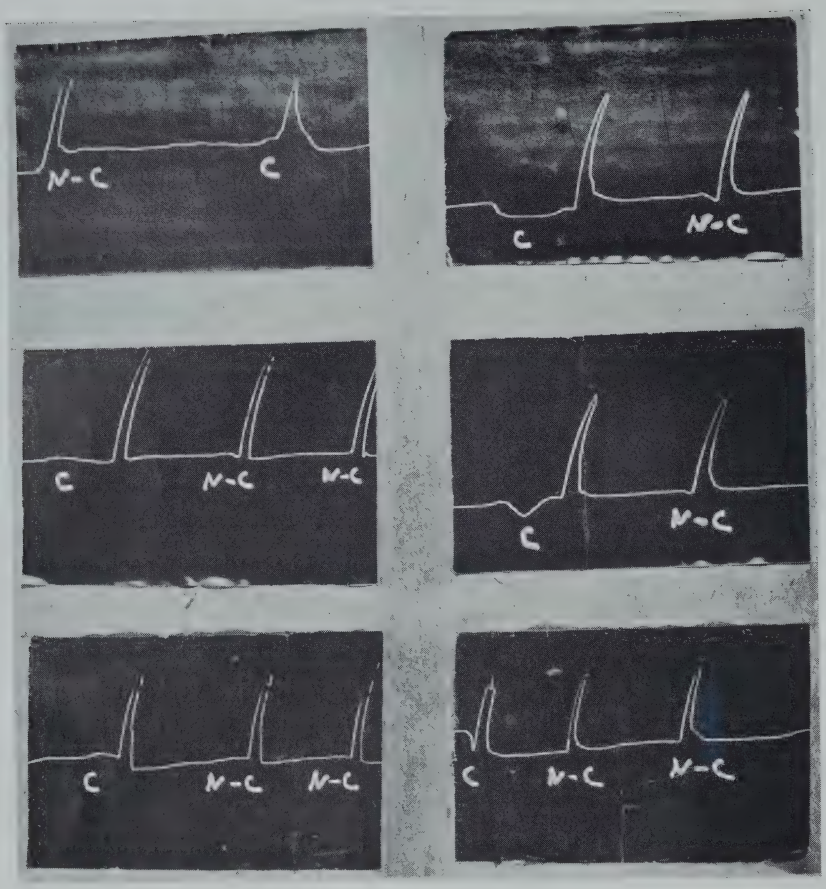


FIGURE 1

response line, and the left-hand record have been deleted in order to conserve space. In no case did the record obtained from the left hand show any appreciable deviation. While Luria obtained marked changes in the left-hand record, this may have been due to the more intense affective situations which he used. In the figure, *C* and *N-C* refer to the responses to the C and N-C words. In all *S*'s the form of the responses to the N-C words was practically the same. We did not find the variety of types of responses which were ob-

tained by Burtt. The reactions of a critical nature were generally in the form of wave-like tremors immediately preceding the motor response. It may be seen in Figure 1 that the differences between the C responses and the N-C responses are for the most part clear cut.

V. Summary and Conclusions

The procedure used in the present study included a preliminary experimental series and a main experimental series. The preliminary series, in which five *S*'s were used, served as a control experiment and helped to standardize the procedure for the main experiment. The main experiment, in which eight *S*'s were used, showed that the Laboratory Deceptive Situation had a definite influence on the later verbal and motor responses. In the case of every *S* the mean and median reaction times to the C words were greater than the mean and median reaction times to the N-C words. The motor responses to the C words were clearly differentiated from the motor responses to the N-C words.

Our results tend to confirm Luria's general thesis that affective arousals in a deceptive situation tend to change the reactive processes of a voluntary behavior system. The affective arousal was specific to the C stimulus arousing it, and did not spread to the N-C words that followed the C words.

Luria's motor method combined with the word-association procedure appears to be a satisfactory method in the study of deception, and it seems that further research and refinement of this technique should yield valuable results on affective processes and on the detection of guilt. Huston, Shakow, and Erickson (9) list a number of problems concerning the affective processes involved in conflicts that may be studied by the use of Luria's technique.

References

1. Benussi, V. Die Atmungsymptome der Luge. *Arch. f. d. ges. Psychol.*, 1914, **31**, 244-273.
2. Blatz, W. E. The cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *J. Exper. Psychol.*, 1925, **8**, 109-132.
3. Brunswick, D. The effects of emotional stimuli on the gastro-intestinal tone. *J. Comp. Psychol.*, 1924, **4**, 19-80.

4. BURTT, H. E. The inspiration-expiration ratio during truth and falsehood. *J. Exper. Psychol.*, 1921, **4**, 1-23.
5. CANNON, W. B. Bodily changes in pain, hunger, fear and rage. New York & London: Appleton, 1915. Pp. xiii+311.
6. CRANE, H. W. A study in association reaction and reaction time. *Psychol. Rev., Monog. Suppl.*, 1915, **18**, No. 80. Pp. 77.
7. CROSLAND, H. R. The psychological method of word association and reaction time as tests of deception. *Univ. Oregon Publ.*, 1929, No. 1. Pp. 104.
8. GASKILL, H. V. The objective measurement of emotional reactions. *Genet. Psychol. Monog.*, 1933, **14**, No. 3. Pp. 177-281.
9. HUSTON, P., SHAKOW, D., & ERICKSON, M. A study of hypnotically induced complexes by means of the Luria technique. *J. Gen. Psychol.*, 1934, **11**, 65-97.
10. LANDIS, C. Studies of emotional reaction. *J. Exper. Psychol.*, 1924, **7**, 325-341.
11. LARSON, J. A. Lying and its detection. Chicago: Univ. Chicago Press, 1932. Pp. xii+453.
12. MARSTON, W. M. Systolic blood pressure symptoms of deception. *J. Exper. Psychol.*, 1923, **6**, 117-163.
13. LURIA, A. R. The nature of human conflicts. (Trans. by W. H. Gantt). New York: Liveright, 1932. Pp. xvii+431.
14. STARR, H. The hydrogen ion concentration of mixed saliva considered as an index of fatigue and emotional excitation and applied to the metabolic etiology of stammering. *Amer. J. Psychol.*, 1923, **33**, 394-418.
15. YERKES, R. M., & BERRY, C. S. The association reaction method of diagnosis. *Amer. J. Psychol.*, 1909, **20**, 22-37.
16. WERTHEIMER, M., & KLEIN, J. Psychologische Tatbestanddiagnostik. *Arch. f. Krim. Anthropol. u. Kriminal.*, 1904, **15**, 72-113.

University of Wisconsin
Madison, Wisconsin

LA MÉTHODE MOTRICE DE LURIA ET L'ASSOCIATION DES MOTS DANS L'ÉTUDE DE LA DÉCEPTION

(Résumé)

La méthode motrice de Luria combinée avec la technique de l'association contrôlée des mots a été employée pour noter les changements dans le système réactif volontaire causé par une situation de déception dans le laboratoire. On a employé les réponses de la main droite pour mesurer les changements moteurs.

On s'est servi de deux séries expérimentales: (1) une série de contrôle où l'on a employé 5 S's, et (2) les principales séries où l'on a employé 8 S's. Dans la série de contrôle les S's ont subi un test préliminaire où les réponses motrices et verbales à une liste de 100 mots servant de stimuli

ont été notées. Vingt de ces 100 mots ont été nommés des mots "critiques"; ces mots ont eu une connotation critique après la situation de déception dans le laboratoire. Après le test préliminaire les S's ont subi une situation de déception dans le laboratoire faite pour faire venir un ton affectif à certaines situations qui seraient rappelées plus tard par les 20 mots critiques. On a fait le dernier test de la même manière que le test préliminaire. Dans la principale série d'expériences on a abandonné le test préliminaire et l'on a réduit le nombre de mots dans la liste d'association de 100 à 35, dont 10 ont été critiques. Les résultats tendent à confirmer la thèse de Luria qu'un réveil affectif dans une situation de déception change les processus réactifs d'un système de comportement volontaire.

RUNKEL

DIE MOTORMETHODE VON LURIA UND WORTASSOZIATION BEI DER UNTERSUCHUNG DES BETRUGS

(Referat)

Lurias Motormethode, welche mit der kontrollierten Wortassoziationsmethode verbunden wurde, wurde zur Beobachtung der Änderungen des willkürlichen reaktiven Systems verwandt, das durch eine Laboratoriumsituation für Betrug verursacht wurde. Die Motoränderungen wurden vermittelst der Reaktionen der rechten Hand gemessen.

Zwei experimentelle Reihen wurden verwandt: (1) eine Kontrollreihe, in der 5 Vpn. gebraucht wurden, und (2) die Hauptreihe, in der 8 Vpn. gebraucht wurden. In der Kontrollreihe erhielten de Vpn. eine Vorprüfung, in der die Motor- und Wortreaktionen auf eine Liste von 100 Reizwörtern bemerkt wurden. Zwanzig dieser 100 Wörter wurden als entscheidene Wörter bezeichnet; diese Wörter hatten entscheidene Mitbedeutung nach der Laboratoriumsituation. Nach der Vorprüfung wurden die Vpn. in eine Betrugsituation im Laboratorium gebracht, die zur Hervorrufung eines affektiven Tones bei gewissen Situationen entworfen wurde, die später durch die 20 entscheidene Wörter hervorgerufen wurden. Der letzte Test wurde auf dieselbe Weise wie bei der Vorprüfung ausgeführt. In der Hauptreihe der Experimente wurde die Vorprüfung aufgegeben und die Anzahl der Wörter in der Assoziationsliste wurde von 100 auf 35 reduziert, deren 10 entscheidene waren. Die Ergebnisse neigen zur Bestätigung Lurias Behauptung, dass die Erzeugung der Affekte in einer Betrugsituation die Rückwirkungsprozesse eines willkürlichen Verhaltenssystems verändert.

RUNKEL

THE SIGNIFICANCE OF THE FLUCTUATIONS EXPERIENCED IN OBSERVING AMBIGUOUS FIGURES AND IN BINOCULAR RIVALRY*

From the Department of Psychology, Duke University

R. Wilfred George

As early as 1906, William McDougall observed (3) that the rate of fluctuation experienced in observing various ambiguous figures was fairly constant for each individual but varied from one individual to another, and suggested that these variations probably indicated some important difference of cerebral constitution. He also observed that the alternation of two differently colored fields, presented simultaneously to the two eyes, is similar in many ways to these fluctuations of ambiguous figures. Later, on the basis of results obtained from a limited number of subjects, he found reason to believe that some correlation obtains between the rates of fluctuation of ambiguous figures and the position of the personalities concerned in the scale of introversion-extroversion (4). He has also shown (5) that changes are brought about in these rates of fluctuation by the administration of certain drugs, the rate being slowed down when alcohol, ether, or other anaesthetic drugs are given to the subject and speeded up when caffein and other of the stimulating alkaloids are given. The first of these two classes of drugs he called the intoxicating or extroverting drugs and the second the introverting or dream producing drugs. On the basis of these and other considerations he developed his application of the chemical theory of temperament to introversion-extroversion.[1] Some of McDougall's

*Accepted for publication by William McDougall of the Editorial Board and received in the Editorial Office, April 22, 1935.

[1]According to McDougall's theory the extroverting drugs act directly upon synapses, raising their resistance, and thus produce a condition of relative dissociation in the brain cortex which results in the freeing of the lower centers (particularly emotional centers in the thalamus) from the inhibiting influence of the higher centers; hence the more freely expressive or more extroverted behavior of the subject under the influence of alcohol, etc. At the same time the extroverting drug slows up the rate of fluctuation experienced in reversible perspective. In the typical extrovert the be-

observations have been called into question by recent experimental results especially those obtained by Guilford (2) and his colleagues at the University of Nebraska. Their results call for further investigation of the problems involved. It was this situation which prompted the present investigation.

The material presented in the following article will have special reference (*a*) to the significance of the fluctuations (experienced in observing ambiguous figures and in binocular rivalry) as a measure of introversion-extroversion,[2] (*b*) to the effect of the stimulating alkaloid and anaesthetic drugs on these same fluctuations, (*c*) to certain sex differences in the rate of fluctuation, (*d*) to certain interesting differences which seem to exist between the rates of fluctuation experienced in binocular rivalry and in reversible perspective.

The subjects used in this investigation were students in the School of Arts and Sciences at Duke University. The first 17 subjects were young men in the Introductory Course in Psychology who volunteered for the experiment. The results obtained from these 17 subjects, as first interpreted, suggested the advisability of selecting further subjects of distinctly introvert or extrovert tendencies. Accordingly only those who were marked extrovert or introvert as measured by a rating scale devised for the purpose of measuring introversion-extroversion (to be described later) were chosen as further subjects. In this manner 22 men were selected from a total number of 99 to whom the rating scale was applied. Nineteen women were selected from a total of 86. According to this criterion 12 of these men were extreme introverts and ten of them were extreme extroverts (two extroverts had trouble with fusion in binocular rivalry so that their scores were omitted in the treatment of the data); ten of the women were extreme introverts and nine of them extreme extroverts. Twenty-four of the men participated in that part of the experiment involving the use of drugs.

---

havior shown and the slow rate of reversal experienced is said to be the result of an extroverting hormone which has somewhat the same effect as the extroverting drug.

[2]Throughout this article the terms extroversion and introversion are used in the modified meaning given them by McDougall in his highly simplified theory which identifies extreme extrovert behavior with uncontrolled expression of all the emotional excitations of the thalamic level and extreme introvert behavior with excessive inhibition of all such immediate expressions, by cortical activity.

Three separate tests involving reversible perspective were employed in measuring the rate of fluctuation. McDougall's windmill (the shadow of the rotating wheel obliquely projected on a screen was observed rather than the wheel itself), Rubin's black and white cross, and a modification of Jastrow's cube illusion. In addition two tests of binocular rivalry were used. The first test involved the color-rivalry of red and green squares and the second the rivalry of diagonal lines running in opposite directions. Subjects were instructed to be passive so far as the rate of fluctuation was concerned for all but the test with McDougall's windmill. On this test they were asked to take in turn three attitudes: a passive attitude, an attitude of speeding up the rate of fluctuations as much as possible, and an attitude of inhibiting the rate of fluctuations as much as possible. This gives us seven tests of fluctuation in all if we consider the three attitudes on the windmill as three separate tests.[3]

In order to determine whether the scores made on tests of fluctuation are correlated with positions in the scale of introversion-extroversion, it was necessary to secure some suitable measure of the latter. A study, however, of the various rating scales that have been published suggested that at least four different traits were being confused under the one heading. The scales included questions which seemed to have reference primarily to self-assertion-submission; others which seemed to distinguish the person who was

[3]Subjects were made thoroughly familiar with each test of fluctuation before any records were taken. Standardized spatial relationships, lighting conditions, and instructions for observing the phenomena and making responses were maintained throughout the experiment. In each test of reversible perspective the subjects sat approximately four meters from the screen on which the phenomenon to be observed appeared. They were asked to fixate a certain point in each figure and to indicate each change observed by pressing a key at their side. In the tests involving binocular rivalry the subjects were first asked to adjust the card so as to obtain the proper focus. It was then kept in that position throughout the remainder of their observations. The order for presentation of the seven tests was as follows: test one (passive attitude on windmill) three one-minute periods; test two (speed up fluctuations on windmill) two one-minute periods; test three (fluctuations on windmill inhibited) four one-minute periods; test two (repeated) two one-minute periods; test one (repeated) three one-minute periods; test four (cube) three one-minute periods; test five (cross) three one-minute periods; test six (diagonals in binocular rivalry) three one-minute periods; test seven (colors) three one-half-minute periods; regular rest periods were introduced between each two successive tests.

given to reflection from the person who was adept at manipulation; others which seemed to distinguish the altruist from the egotist; and finally others which seemed to distinguish those whose behavior is habitually restrained or inhibited from those who are relatively unrestrained or freely expressive.

In order to check this analysis and to obtain a more adequate rating scale for our purpose some preliminary work was necessary. A list of 107 questions was developed with the four above-mentioned distinctions in mind. Most of the questions were taken directly or in modified form from current rating scales for introversion-extroversion and from Allport and Allport's (1) test of ascendency and submission. A few others were added which were thought to have significance for the traits in question. Since we were especially interested in developing a rating scale which would meet the requirements of a particular concept of introversion-extroversion, it was desirable to obtain judgments on this preliminary rating scale by those individuals who were most familiar with this concept. Accordingly Dr. McDougall and two of his associates, Drs. Lundholm and Zener, were asked to check the rating scale. The questions were arranged in random order and presented with appropriate instructions.[4] The questions had already been classified by the writer before they were arranged in random order for checking so that we have four independent judgments on each question.

The final rating scale resulting from this preliminary work includes questions relative to all four sets of traits. Questions for egotism-altruism, manipulation-reflection, and self-assertion-submission were included, providing that when they were checked on the preliminary rating scale at least three judges agreed as to their diagnostic value for the trait in question and for no other and providing the fourth judge did not check the question at all. Twenty-one out of a total of 27 self-assertion-submission questions met this standard; nine out of a total of 20 manipulation-reflection ques-

[4]The judges were asked to check through the entire list of questions four times. The first time they were to determine whether each question was suitable for distinguishing introverts from extroverts. The other three complete checks were to determine the suitability of each question for distinguishing each of the other three sets of traits. Each question could be checked as many times as it seemed necessary to differentiate the trait in question.

tions; and 17 out of a total of 33 egotism-altruism questions. (Eighteen out of a total of 35 introvert-extrovert questions conformed to this standard also.) The introvert-extrovert questions included in the final rating scale were determined by a further check. Certain additional items were added to the original lists, making 45 in all. Each question was again checked by the judges mentioned above. After each question the numbers 1, 2, and 3 appeared. If the question was one which was judged well suited to differentiate between an extrovert and an introvert, number 1 was checked; if the question was one which might be expected to differentiate between an extrovert and an introvert, but somewhat less clearly, number 2; if the question was one which did not seem to differentiate between an extrovert and an introvert or was one in which its use for this purpose is questionable, number 3 was checked. Questions from this list were finally considered diagnostic for introversion-extroversion and included in the final rating scale providing (1) the average for the numbers checked was not greater than two and (2) no one judge checked three, thus indicating that in his opinion the question is not a good one for the purpose. Thirty-eight of the items met this standard. Seventeen of these questions when answered in the affirmative indicate introversion and 21 of them when answered in the affirmative indicate extroversion. These 38 questions, together with the questions previously obtained for egotism and altruism, reflection and manipulation, self-assertion and submission, constituted the main body of the final rating scale.[5]

Although there was not complete agreement as to the significance

[5]After each question in the final rating scale there was provided a scale of five equal steps running from 0, standing for a distinctly negative answer, to 4, standing for a distinctly positive answer. The introvert-extrovert score for each subject was obtained by subtracting the sum of the numbers checked for extrovert questions (questions which when answered affirmatively indicate extroversion) from the sum of the numbers checked for introvert questions; likewise scores for egotism-altruism, manipulation-reflection and self-assertion-submission. This method of determining scores is in line with current treatment. Allport speaks of ascendency and submission as two separate traits but says they may be treated statistically with reference to one scale. McDougall says the marked introvert and the marked extrovert determine the two ends of a single linear scale, the various positions on which are due to the degree of intensity of a single personality factor. Lundholm points out that when egotism is strong then the capacity to feel with and for others disappears and vice versa.

of the questions contained in our preliminary rating scale for measurement of personality traits, yet the agreement is sufficiently close to lend some weight to the analysis of other rating scales for introversion-extroversion suggested above. The value of making the distinctions mentioned is further indicated by a comparison of the scores obtained by our group of 56 subjects for each set of traits distinguished. (Manipulation-reflection is not included in this comparison because of the very few items concerning which there was agreement as to their diagnostic value for this trait.) In Table 1 is

TABLE 1

	1	2	3	4
1. Self-assertion-submission		.26	—.35	—.06
2. Egotism-altruism	.26		.10	.26
3. Introversion-extroversion	—.35	.10		.48
4. Rubin's cross	—.06	.26	.48	

presented the intercorrelation of the scores for these traits, together with the scores on Rubin's cross test (the test of fluctuation which gave the highest correlation with the introvert-extrovert rating scale).

The correlation coefficient is less than three times its probable error for all but two of the correlations in this table. The correlation between the scores on the introvert-extrovert rating scale and the rate of fluctuation on the cross is approximately seven times its probable error, while that between the scores for introversion-extroversion and for self-assertion-submission is approximately three times its probable error. These results emphasize the necessity of drawing clear distinctions if we are ever to get a satisfactory measure of any one or all of these traits. It was with this point clearly in mind that each of the judges checked the questions from which our rating scale for introversion-extroversion was chosen. This rating scale should therefore be a relatively valid measure of degrees of introversion-extroversion.[6]

The validity of a rating scale is sometimes said to be determined

[6]Anyone interested in the rating scale devised for use in this investigation may obtain a copy by sending a request for the same to the author in care of the Department of Psychology, Tarkio College, Tarkio, Mo.

by the degree to which self-ratings and the ratings of close friends agree. A separate rating scale including all introvert-extrovert questions on which the subjects rated themselves was devised for the purpose of securing the judgment of close friends as to the prevalence of the trait in question in each subject. The average scores thus obtained from the judgment of three friends of each of our 56 subjects were correlated with the scores obtained when the subjects judged themselves according to the same criterion. The coefficient of correlation was .68.

The validity of the tests of fluctuation is fairly well assured by the technique for administering the tests.

The reliability of each test of fluctuation was measured by correlating the scores for 25 subjects on one day with the scores of the same 25 subjects on a second day. The reliability of the introvert-extrovert rating scale was determined by correlating the scores on one half with the scores on the other half of the rating scale. Each half of the extrovert questions was selected by taking every other extrovert question throughout the rating scale; likewise the introvert questions. The reliability of the short scales developed for egotism-altruism and self-assertion-submission was determined in the same manner, except that the scores for the two halves of the scale were obtained by an arbitrary division of the scale rather than by taking every other item, due to the manner in which these scales were graded. The scores of the 56 individuals who were also subjects for the test of fluctuation were those used in computing the reliability of the rating scales. This procedure gives the reliability of each test as follows.

Windmill—passive attitude	.86±.02
Windmill—speeding up fluctuations	.64±.05
Windmill—inhibiting fluctuations	.74±.04
Cube	.83±.03
Cross	.83±.03
Binocular rivalry—diagonals	.78±.04
Binocular rivalry—colors	.74±.04
Introvert-extrovert rating scale	.74±.04
Egotism-altruism rating scale	.72±.04
Self-assertion-submission rating scale	.60±.06

The coefficients of reliability for the two active attitudes on the windmill are rather low, as might be expected because of the possible difference in degree of effort exerted at different times. The attitude

of speeding up the reversals is further complicated by the fact that certain cues were often discovered even after considerable experience in trying to make the reversals take place rapidly. It was impossible to overcome this irregularity by preliminary practice since sometimes it would appear quite late in the experiment. It did obviously tend to reduce the reliability of the measure. In less degree perhaps the same could be said for the attitude of inhibiting the rate of fluctuation. Considering the nature of the remainder of the tests, the reliabilities may be said to be average or better.

Comparison of Scores on Tests of Fluctuation with Scores on the Final Rating Scale

In Table 2 we have a summary of the intercorrelations among the various tests of fluctuation and the rating scale. In addition, line *"A"* gives the correlations corrected for attenuation between the rating scale and each of the tests opposite on the left.

With very few exceptions the correlations in question are high enough to be significant. They seem to indicate that there is a considerable relationship among the various tests of fluctuation as well

TABLE 2

Intercorrelations of Tests of Fluctuation and the Rating Scale

	1	2	3	4	5	6	7	8	*A*
1. Windmill—passive		64	74	85	70	53	55	44	56
2. Windmill—fluctuations speeded up	64		51	24	56	38	49	31	45
3. Windmill—fluctuations inhibited	74	51		35	42	22	25	32	43
4. Cube—passive	85	24	35		72	50	53	36	46
5. Cross—passive	70	56	42	72		54	49	48	62
6. Binocular rivalry—passive (diagonals)	53	38	22	50	54		73	37	49
7. Binocular rivalry—passive (colors)	55	49	25	53	49	73		33	45
8. Rating scale	44	31	32	36	48	37	33		

as between the various tests of fluctuation and the rating scale. Spearman's method of tetrad differences for determining the presence of a common factor influencing scores in a group of tests, given the intercorrelations among those tests, offers itself as a convenient means of determining if the relationship indicated by our intercorrelations is due to the presence of such a common factor. According to the theory, if the tetrad differences based upon a group of intercorrelations approximate to zero within the limits of chance variation, then one common factor is present in the variables involved, while all other factors are specific to each particular function. If there is overlapping of the specific factors, then the criterion is not satisfied and the tetrads fall proportionately farther away from zero. Spearman (7, Appendix, p. xi) gives a formula for determining the probable error of the tetrads. Given the probable error and our distribution of tetrad differences, it is a simple matter to compare the number of tetrads which actually fall within the probable error with the number which would be expected to fall within the probable error by chance. If the former figure approximates the latter, then the criterion has been satisfied and a single factor is being measured by

TABLE 3

EMPIRICAL AND EXPECTED DISTRIBUTION OF ALL THE TETRADS

	Empirical	Expected
1 P.E.	70	105
2 P.E.	56	69
3 P.E.	29	27
4 P.E.	22	8
Beyond 4 P.E.	33	1
	$P.E._t$ .0456	Median .0730

TABLE 4

EMPIRICAL AND EXPECTED DISTRIBUTION OF TETRADS INVOLVING TESTS 1, 2, 4, 5, 6, AND THE RATING SCALE

	Empirical	Expected
1 P.E.	20	23
2 P.E.	13	14
3 P.E.	6	6
4 P.E.	4	2
Beyond 4 P.E.	2	0
	$P.E._t$ .0516	Median .0693

TABLE 5

EMPIRICAL AND EXPECTED DISTRIBUTION OF TETRADS INVOLVING TESTS 1, 4, 5, 6, AND THE RATING SCALE

	Empirical	Expected
1 P.E.	7	8
2 P.E.	4	4
3 P.E.	3	2
4 P.E.	1	1
Beyond 4 P.E.	0	0
	P.E.$_t$ .0465	Median .0648

our battery of tests. If, on the other hand, many of the tetrads fall beyond the point on the distribution curve where they would be expected to fall by chance, then the criterion has not been satisfied. Tables 3, 4, and 5 show the empirical and expected frequency of the tetrad differences derived from the intercorrelations of the tests indicated above.

From the empirical and expected distribution of all the tetrads as presented in Table 3, it is obvious that they do not satisfy the criterion. There is some *a priori* basis for believing that, in addition to a possible common factor present in the whole group of tests, we have a specific factor common to tests two and three (rates of fluctuation on windmill speeded up and rates of fluctuation on windmill inhibited) and perhaps all three windmill tests, and also another specific factor common to tests six and seven (the two tests of binocular rivalry). Of the 33 tetrads which fall beyond the fourth probable error as indicated in Table 3, 14 involve both tests six and seven, and eight of them involve both tests two and three. All other tetrads falling beyond the fourth probable error, with one exception, involve some combination of tests one, two, and three (windmill tests). Since if the criterion were satisfied all of these except one should fall within the fourth probable error, we have sufficient evidence that there is a special factor common to each of the two pairs of tests. This is probably because each pair of tests measures the same factor. When in Table 4 we have the empirical distribution and the expected distribution of the tetrads involving tests 1, 2, 4, 5, 6, and the rating scale, then we find that the empirical distribution approximates the expected distribution, indicating that the criterion has been satisfied and that therefore we probably have a common factor run-

ning throughout these six tests and that no specific factor is common to any two of them. The two tetrads falling beyond the fourth probable error involve tests 1 and 2. The criterion is found to be even better satisfied if we compare the empirical and expected distribution of tetrads involving tests 1, 4, 5, 6, and the rating scale as indicated by Table 5.

Spearman (7, p. xvi) also supplies us with another formula whereby we can determine how much of the common factor is present in each of our measures. Applying this formula to our data we obtain the following correlations between the common factor and the various tests of fluctuation and between the common factor and the rating scale.

Windmill—passive attitude	.96
Windmill—fluctuations speeded up	.63
Windmill—fluctuations inhibited	.55
Cube—passive	.73
Cross—passive	.82
Binocular rivalry—diagonals (passive)	.66
Binocular rivalry—colors (passive)	.69
Rating scale	.52

These correlations indicated a rather marked presence of a common factor throughout these tests. The similarity of the correlations for tests 2 and 3 is usually interpreted to indicate that the two are measuring the same specific factor. This seems quite probable in view of the fact that they both involve a voluntary effort to modify the rate of reversal. The similarity of tests 6 and 7 is even more marked. These results are in line with the discussion of the distribution of tetrad differences given above.

In the light of the above results there seems to be a definite relationship between scores made on the tests of fluctuation and scores made on the introvert-extrovert rating scale. It is in place here to make suggestions as to certain factors which have tended to obscure this relationship in other investigations. One suggestion we have already implied, namely, that the scales developed for the purpose of measuring introversion-extroversion have been too inclusive; they have not related to one trait but to many traits, a fact which is indicated by the lack of correlation between the various rating scales themselves, as well as the material presented in connection with the derivation of the rating scale used in this investigation.

A second suggestion, which comes partly from clinical observation, concerns the difficulty which the average subject has in distinguishing one trait from another. The failure to distinguish self-assertion from extroversion furnishes a good example of this difficulty. Here we call attention to the correlation of —.35 between self-assertion-submission and introversion-extroversion. If this relationship be due to a failure to distinguish one of these traits from the other such a failure would obviously affect relationships between scores on the introvert-extrovert rating scale and any other test for introversion-extroversion which might be used.

One other suggestion we wish to make as to a possible factor which has tended to affect the relationship in question. This suggestion is that education and environmental influences in general may play an important part in determining responses made on an introvert-extrovert rating scale, while such influences probably do not appreciably affect the performances on our tests of fluctuation. If the effect of this environment is to modify scores on the rating scale as much in one direction as in the other, then it becomes very difficult to get any measure which will indicate this effect. If, however, the effect of training and environmental influences in general tend in one direction more than in the other, then it should be possible to get a measure which will show that such a tendency exists.

If the differential effect of the environment suggested above is predominately in the direction of raising the extrovert score of subjects, then, if we wish to secure subjects as little influenced by this differential factor as possible, we could best choose our introverts according to the rating scale, for they would constitute the group which is relatively free from the extroverting influence. The extroverts, on the other hand, we could best choose according to the test of fluctuation, lest by choosing them according to the rating scale we obtain those who were extroverted because they had been more than usually influenced by the extroverting environment. But if the differential effect of the environment is predominately introverting in nature and we wished still to secure subjects as little influenced by this differential factor as possible, we could best choose our subjects in the reverse fashion. Accordingly 14 extreme introverts and 14 extreme extroverts were chosen according to each of the two methods suggested above. Table 6 gives the results of

TABLE 6

CORRELATIONS BETWEEN SCORES ON TESTS OF FLUCTUATION AND SCORES ON RATING OBTAINED FROM VARIOUSLY SELECTED GROUPS

	A	*B*	*C*
Windmill—passive	82	58	44
Windmill—fluctuations speeded up	70	65	31
Windmill—fluctuations inhibited	65	59	32
Cube—passive	75	63	36
Cross—passive	80	64	48
Binocular rivalry—passive (diagonals)	71	56	37
Binocular rivalry—passive (colors)	79	66	33
Average	75	62	37

correlating the scores on tests of fluctuation with scores on the rating scale for each group of 28 individuals. In line *A* we have the correlation obtained when the selection was made on the assumption of a predominately introverting environment, in line *B* the correlations obtained when the selection was made on the assumption of a predominately extroverting environment, and in line *C* the correlations for the entire group of 56 subjects with no selection made at all. It is of course true that the correlations in lines *A* and *B* of our table are spuriously high on account of our selection. The significant thing about them, however, is the rather marked difference between the correlations in lines *A* and *B*.[7] This result would suggest a predominately introverting environment influencing differentially the scores made on the rating scale as compared with scores on the tests of fluctuation.

It is possible that the tests of fluctuation measure more accurately at the higher end of the scale than at the lower. Two things, at least, would seem to indicate that this is not the case. In choosing the 14 individuals who had the highest and lowest rates of reversal for each test respectively, a smaller total number was necessary

[7]It is interesting to notice that there was an average overlap of only 1.5 subjects per test when choice was made according to the first criterion mentioned above and of less than one subject per test when the second criterion was used. By overlapping we mean that a subject who was chosen among the first 14 also fell among the second 14 for the selection in question. In such a case he was omitted from the second 14 and the next subject in line was used in his stead. This lack of overlapping itself is significant for the relationship between introversion-extroversion and the tests of fluctuation.

for those who had low scores than those who had high scores, indicating that those who had low scores on one test were a little more likely to have low scores on the others, than were those who had high scores on one test likely to have high scores on the others. Furthermore, an inspection of the distribution shows rather clearly that it is not the individuals who fall at the lower end of the scale measuring rate of fluctuation who make the difference but rather those who tend more toward the middle. There are, no doubt, other interpretations, but we believe that the one offered above is worth considering. We suggest, therefore, the desirability of distinguishing two scales upon which individuals may be placed: an introvert-extrovert scale and what we will call an *I-E* scale. The position which one holds on the first will be determined by the score on the rating scale. The position on the *I-E* scale will be expressed in terms of the scores on the tests of fluctuation (subjects on the *I* end of the scale have high rates of fluctuation; those on the *E* end of the scale have low rates of fluctuation). Our clinical data, omitted on account of space limitations, emphasize the basis on which this distinction is made. It was these data in fact which first led to the brief analysis of scores presented above in connection with this suggestion. It will be noted that both this suggestion and the one previous to it have a bearing on results reported in this study.

The possibility that tradition and custom may profoundly modify an otherwise extroverted people, so that their behavior is in many respects characteristically introverted, is suggested in an article on Japanese temperament and character by Seligman (6). Seligman believes the Japanese temperament to be typically extroverted and at the same time he finds a sort of overcompensation for this extroverted characteristic of their behavior in the customs and traditions of the people. Is it not possible that a somewhat similar influence in our own country may have had a like effect and may be responsible for the differential effect which we have suggested?

Drug Effects

Sodium amytal and caffein were used in our study of the effects of anaesthetics and of stimulating alkaloids on the phenomena in question. Five grains of sodium amytal and nine grains of caffein citrate were used as doses. The doses of both drugs were admin-

istered orally in capsules which looked exactly alike. The smaller amount of the sodium amytal was compensated for by the addition of a little milk sugar in the capsule. This made possible a certain control over the factor of suggestion. Furthermore, precautions were taken against leading the subjects to expect any particular effects from the taking of the drugs.

Twenty-two of our 24 subjects were given amytal and 15 of them were given caffein. The experimental procedure was as follows—each subject was first tested on all tests in exactly the same way as all other subjects were tested, after which each was given the drug. After they were given sodium amytal 55 minutes were allowed to elapse for the drug to take effect and they were again tested in the same manner. After they were given caffein, one hour and 15 minutes were allowed to elapse before the second series of tests was given. It is important that the drug be given sufficient time to take effect. In preliminary experiments the failure to allow a sufficient length of time had given ambiguous results. There was some evidence that the time intervals just indicated above were too short to obtain the maximum drug effect. This is consistent with McDougall's (5) finding that caffein does not attain its maximum effect until some two hours after administration.

We determined for each individual the percentage of increase above or decrease below the normal rate of fluctuation which resulted from taking our series of tests following administration of each drug. The average percentage for those taking sodium amytal shows a decrease in rate of fluctuation on all tests as follows: windmill (passive) 26 per cent; windmill (fluctuations speeded up) 21 per cent; windmill (fluctuations inhibited) 26 per cent; cube 23 per cent; cross 16 per cent; binocular rivalry (diagonals) 16 per cent; binocular rivalry (colors) 16 per cent. The average percentage for those taking caffein shows an increase in rate for all tests as follows: windmill (passive) 36 per cent; windmill (fluctuations speeded up) 31 per cent; windmill (fluctuations inhibited) 55 per cent; cube 25 per cent; cross 36 per cent; binocular rivalry (colors) 16 per cent; binocular rivalry (diagonals) 13 per cent. Contrary to the general trend, in the tests of reversible perspective, an average of 2.6 persons per test showed an increase in rate of fluctuation when sodium amytal was given and an average of 3.6

persons per test showed a decrease when caffein was given. Likewise in the tests of binocular rivalry an average of five persons per test showed an increase when sodium amytal was given and an average of four persons per test showed a decrease when caffein was given. In most cases these deviations from the general trend were small.

Data on the relative effects of drugs on scores of individuals who fall at either end of the *I-E* scale, while not clear cut, suggest that those at the *"I"* end of the scale are slightly more influenced by sodium amytal than are those at the *"E"* end and that the reverse is is even more noticeably true of the effect of caffein.

Even though we recognize the various sources of error in drug experiments only partially controlled, still it may fairly be said that the general trend of the results given above is clear, namely, that sodium amytal tends to reduce the number of fluctuations on our tests and that caffein on the other hand tends to increase the number of these fluctuations. This is in keeping with McDougall's observation.

Sex Differences

Although our subjects are too few to warrant drawing any final conclusions as to sex differences, yet the results were sufficiently interesting to justify reporting them.

It will be remembered that the principal subjects for this study included 19 women and 37 men. Seventeen of the men were a volunteer group from a class in general psychology; the other 20 men and the 19 women were selected because they were either extreme extroverts or extreme introverts as determined by scores on the introvert-extrovert rating scale. According to this criterion 12 of these men were introverts and eight of them were extroverts; ten of the women were introverts and nine of them were extroverts.

Table 7 presents a summary of the scores for men and women on the seven tests of fluctuation and on the rating scale. In column *A* we have the average score for women on each of the eight tests and in column *B* the range of scores for women on each test in columns *C* and *D* the average scores and range on each test for men selected on the basis of rating scale as indicated above; in columns *E* and *F* the average score and range on each test for the group of volunteer men.

TABLE 7

SEX DIFFERENCES ON TESTS OF FLUCTUATION AND THE RATING SCALE

	A	*B*	*C*	*D*	*E*	*F*
Windmill—passive	7.6	3.2 to 15.8	12.4	3.7 to 29.7	8.4	2.5 to 18.5
Windmill—fluctuations speeded up	11.8	4.5 to 21.8	27.3	8.8 to 54.3	14.5	4.8 to 39.0
Windmill—fluctuations inhibited	5.3	1.8 to 10.3	7.7	2.8 to 15.5	5.8	1.0 to 11.0
Cube—passive	10.1	3.7 to 21.7	15.6	7.3 to 43.7	11.8	4.7 to 24.3
Cross—passive	11.8	2.3 to 20.3	17.0	7.7 to 34.3	12.1	4.7 to 24.3
Binocular rivalry—passive (diagonals)	28.9	11.3 to 21.3	30.9	16.3 to 57.3	25.7	5.3 to 55.0
Binocular rivalry—passive (colors—30 sec.)	13.2	7.0 to 18.7	14.5	8.3 to 23.3	10.8	5.3 to 16.0
Rating scale	—4.6	30 to —40	1.3	38 to —38	—12.9	12 to —38

A number of things should be noticed in connection with Table 7.

1. According to the scores on the rating scale the selected men are somewhat more introverted than are the women, while the volunteer group of men are more extroverted. This is probably due to the fact that among the selected men we have a number of pronounced introverts, while the other group of men who volunteered for the experiment were naturally more extroverted.

2. Throughout the first five tests which involve reversible perspective, the men have distinctly higher scores than the women, indicating that they fall more toward the *I* end of the *I-E* scale. If we take the scores on the cross test as representative, we find an average difference between the selected men and women of 5.2 (range for men 6 to 34 and for women 2 to 20). The standard error of the difference is 1.8, which makes the critical ratio equal to 2.9. Furthermore it is significant that only two women exceeded the median for selected men. It must be remembered, of course,

that this result is based upon the scores of a limited number of men and women who are extreme on the introvert-extrovert scale. The fact, however, that the group of volunteer men also has somewhat higher scores on the tests of fluctuations than the selected women indicates a real difference between the two sexes.

3. A comparison of the scores for selected men and women for the tests of binocular rivalry gives us results which are in rather marked contrast to those just mentioned. The scores of the women for these tests are almost as high as those for the men and are noticeably higher than the scores for the group of volunteer men.[8]

Comparison of Scores on Tests of Binocular Rivalry and Reversible Perspective

We wish to call attention to data in three different connections which have a bearing on the relationship between scores made on the tests of binocular rivalry and scores made on tests of reversible perspective.

1. Results reported on a previous page (53) indicate that there is a differential drug effect when the tests of binocular rivalry are compared with the tests of reversible perspective. The average increase in the rate of fluctuation of the latter when caffein was given was 36.6 per cent; for the former it was 14 per cent. The average decrease of the latter when sodium amytal was given was 22.4 per cent and for the former 16 per cent. These figures, therefore, indicate that the rate of fluctuation in binocular rivalry is con-

---

[8]Preliminary group tests given to junior high school students indicate that the relationship between scores of boys and girls of this age may be just the opposite to that found for men and women of college age. Girls at the junior high school age fall, on the average, noticeably closer to the *"I"* end of the scale than boys of the same age. If this difference be further substantiated, it suggests the possibility that a factor influencing the rate of fluctuation is somehow related to the glandular activity associated with sex development of the individual.

Preliminary group tests, given to colored men and women of college age and rank, indicate that they score, on the average, closer to the *"E"* end of the scale than white men and women of the same age and rank.

These preliminary group tests raise questions as to the significance of size, clearness, and brightness of figure and background in the determination of fluctuation rate.

siderably less affected by the administration of both caffein and sodium amytal than are the scores for reversible perspective.[9]

2. The scores presented in Table 7 suggest that there is not the same sex difference for scores on binocular rivalry tests as was found to exist for scores on tests of reversible perspective. When the scores of the two similar groups of selected men and selected women are compared it is found that on the five tests of reversible perspective the men have scores which are on the average 67.4 per cent higher than are the scores for the women. On the test in which the difference is least, the men have a score which is 44 per cent higher than that of the women. As opposed to this when we compare the scores of the same two groups on the tests for binocular rivalry it is found that the men have scores which are on the average only 8.5 per cent higher than are the scores for women. If we compare the scores of the selected group of women with the scores of the more extroverted group of volunteer men still we find that the scores of these men on the tests of reversible perspective are also higher than are those of the women, but the scores on the tests of binocular rivalry for these men, however, are noticeably lower than are the scores for the women.

3. Space will permit but a brief reference to a third set of facts which are interesting for the distinction we are making. These facts relate to the results of some preliminary experiments pertaining to voluntary control of fluctuations experienced in the two types of tests in question. Tests were made to determine the voluntary control of rates of fluctuation in binocular rivalry and reversible perspective first with the eyes normal, then with the intrinsic muscle of one eye paralyzed, and finally with the intrinsic muscles of both eyes paralyzed. Under normal conditions the voluntary attempt to control the rate of fluctuation in tests of reversible perspective seems to be about twice as effective as in the case of voluntary control of the fluctuations in binocular rivalry. Paralysis of the intrinsic muscles of the eye seemed to have relatively little effect on voluntary control in tests of reversible perspective, but in tests of binocular

[9]In connection with this differential drug effect we call attention to McDougall's finding that alcohol successively weakens and suspends the hierarchy of functions of the brain and therefore of the mind, in the order from above downward (5).

rivalry the control was noticeably decreased. Voluntary control of rivaling after-images also seemed to be noticeably affected when the intrinsic muscles of the eye were paralyzed.

Data reported earlier in this article indicate that the tests of binocular rivalry measure a specific factor which is not measured by the tests of reversible perspective. It is rather generally assumed that perception involved in reversible perspective is a higher level function than that involved in binocular rivalry. The results of our preliminary experiments relative to voluntary control of the two seem to be in keeping with this assumption. We therefore suggest that the specific factor peculiar to the tests of binocular rivalry is related to the lower level function involved and that it is because binocular rivalry is a lower level function that it is less directly under voluntary control and likewise less subject to the influence of drugs. The sex difference which was also found to exist is difficult to interpret but is interesting nevertheless.

Finally, it should be noted that, if it be true that binocular rivalry and reversible perspective are dependent upon functions of different levels, then the fact that scores made on both are related to scores made on our introvert-extrovert rating scale is especially significant. Whatever characteristic it is, then, which is related to introversion-extroversion and which we are measuring with our tests of fluctuation, it is a characteristic which pertains to at least two levels of the nervous system. It might fairly be supposed that its influence is more general.

Conclusions

Within the limits of the number of subjects selected for this study and the particular sort of selection made, the results of this investigation seem to warrant the following conclusions.

1. Scores made on the tests of fluctuation used in this study, including tests of both binocular rivalry and reversible perspective, show significant correlation with scores made on our rating scale for introversion-extroversion. This rating scale was developed for the purpose of rating individuals in respect to freedom and ease of expression, a trait which McDougall regards as a true unitary trait and as the most essential factor in determining any subject's position on the introvert-extrovert scale. The data obtained in the develop-

ment of this rating scale emphasize the importance of defining personality traits as accurately as possible.

2. Our tests of fluctuation measure one common factor; this is true at least when only one windmill test and one test of binocular rivalry are considered. There is probably a specific factor common to the two tests of binocular rivalry and another to the two active attitudes on the windmill.

3. Three suggestions are given as to factors which influence relationship between the scores on the tests of fluctuation and scores on the introvert-extrovert rating scales. Two of them deal with difficulties in securing a proper introvert-extrovert score; the third relates to a differential environmental effect on the two sets of scores in question. Each of these factors apparently tends to minimize the relationship between scores on an introvert-extrovert rating scale and scores on our tests of fluctuation.

4. Sodium amytal reduces the number of fluctuations in tests involving binocular rivalry and reversible perspective, while caffein increases the number of these fluctuations.

5. Women have distinctively lower scores than men on our tests of reversible perspective; the scores for men and women on tests involving binocular rivalry are approximately the same. The results obtained for junior high school boys and girls in preliminary group tests involving reversible perspective gave higher scores for girls than for boys. (The higher the score the more an individual tends toward the *I* end of our *I-E* scale.) This difference between junior high school and college young people is especially interesting in view of the difference in maturation of the two groups.

6. A comparison of scores on tests involving reversible perspective with those involving binocular rivalry indicates the following differences.

a. Sex difference. The scores of the men are distinctly higher on tests of reversible perspective than scores of women. On the other hand the scores of men and women are nearly the same on tests of binocular rivalry.

b. Differences in drug effects. The scores on the tests of reversible perspective are appreciably more affected by both caffein and sodium amytal than are the scores on the tests involving binocular rivalry.

c. Difference in ease of voluntary control. Both when the eyes are normal and when the intrinsic muscles are paralyzed, voluntary control of fluctuations experienced in reversible perspective is appreciably greater than in binocular rivalry. It is most noticeable when the intrinsic muscles of the eyes are paralyzed.

The suggestion is made that the differences found in the comparison of scores made on tests of reversible perspective and binocular rivalry are due to a difference in level of function. If such be the case it has interesting implications for the relationship we found to exist between scores made on various tests of fluctuation and scores made on our rating scale for introversion-extroversion.

References

1. Allport, G. W. A test for ascendance-submission. *J. Abn. & Soc. Psychol.*, 1928, **23**, 118-136.
2. Guilford, J. P., & Braly, K. W. An experimental test of McDougall's theory of extroversion-introversion. *J. Abn. & Soc. Psychol.*, 1931, **25**, 382-389.
3. McDougall, W. The physiological factors of the attention process. *Mind*, 1906, **15**, 329-359.
4. ———. Outline of abnormal psychology. New York: Scribner's, 1926. Pp. xiii+566.
5. McDougall, W., & Smith, M. The effects of alcohol and some other drugs during normal and fatigued conditions. London: H. M. S. Office, 1920.
6. Seligman, C. S. The Japanese temperament and character. *Trans. Japan Soc. London.*, 1931, **28**, 124-138.
7. Spearman, C. The abilities of man. London: Macmillan, 1927. Pp. xxiii+145. (See esp. Appendix, p. xi.)

Tarkio College
Tarkio, Missouri

LA SIGNIFIANCE DES FLUCTUATIONS ÉPROUVÉS DANS L'OBSERVATION DES FIGURES AMBIGUÉS ET DANS LA RIVALITÉ BINOCULAIRE

(Résumé)

Les résultats de cette étude indiquent qu'il y a un facteur commun qui joue un rôle dans la détermination de la vitesse du renversement éprouvé dans l'observation des figures ambiguës et dans la rivalité binoculaire; aussi que ce même facteur est important pour déterminer les résultats faits sur une échelle d'évaluation des introvertis et des extrovertis quand cette

échelle est construite pour mesurer les différences en degrés de la force d'expression, la base primaire sur laquelle le Dr. William McDougall fait la distinction entre un introverti et un extroverti. Les introvertis ont de grandes vitesses et les extrovertis ont de petites vitesses de fluctuation. Les résultats sont basés sur les résultats obtenus de trois tests comprenant des figures ambiguës et deux tests de rivalité binoculaire, avec une échelle d'évaluation appropriée pour l'introversion et l'extroversion.

Les résultats des tests comprenant des figures ambiguës quand on les compare aux résultats des tests comprenant la rivalité binoculaire indiquent (1) que tandis que les fluctuations éprouvées dans les deux sont réduites par des drogues anesthésiques et augmentées par les alcaloïdes stimulants, celles-là sont beaucoup plus affectées que celles-ci; (2) que le contrôle volontaire de celles-là est plus efficace que celles-ci quand les yeux sont normaux et surtout quand les muscles intrinsèques des yeux sont paralysés; (3) qu'une différence appréciable entre les sexes se montre dans les résultats comprenant des figures ambiguës en faveur d'un résultat plus élevé pour les hommes mais que les différences dans les tests comprenant la rivalité binoculaire, bien que dans la même direction, ne sont que très légères.

GEORGE

DIE BEDEUTUNG DER SCHWANKUNGEN, DIE DURCH DIE BETRACHTUNG ZWEIDEUTIGER FIGUREN UND BEI ZWEIÄUGIGEM WEITTEIFER ENTSTEHEN

(Referat)

Die Ergebnisse dieser Untersuchung weisen darauf hin, dass es einen gemeinsamen Faktor gibt, der eine Rolle bei der Bestimmung der Schnelligkeit der Umkehrungen spielt, die durch die Betrachtung zweideutiger Figuren und bei zweiäugigem Weitteifer entstehen; auch dass dieser Faktor zur Bestimmung der Werte wichtig ist, die bei einer Introversion-Extraversionskala erhalten wurden, wenn die Skala so konstruiert ist, dass Unterschiede in Graden der Ausdrücklichkeit gemessen werden, die Hauptbasis darstellt, auf die Dr. William McDougall den Unterschied zwischen dem Introvertierten und dem Extravertierten gründet. Die Introvertierten haben hohe Schnelligkeiten und die Extravertierten haben niedrige Schnelligkeiten der Schwankung. Die Ergebnisse beruhen auf den Werten, die durch drei Tests mit zweideutigen Figuren und zwei Tests des zweiäugigen Weitteifers zusammen mit einer geeigneten Skala für Introversion-Extraversion erhalten wurden.

Wenn die Werte für die Tests mit zweideutigen Figuren mit den Werten für die Tests mit zweiäugigen Weitteifer verglichen werden, so finden wir, (1) dass, obgleich die in beiden erlebten Schwankungen durch betäubende Arzneien vermindert und durch anregende Alkaloide vermehrt werden, die ersteren merklich mehr beeinflusst werden als die letzteren; (2) dass die willkürliche Kontrolle der ersteren wirksamer ist als bei den letzteren, wenn die Augen normal sind und besonders wenn die innerlichen Muskeln der Augen gelähmt werden; (3) dass ein bedeutsamer Geschlechtsunterschied bei Werten mit zweideutigen Figuren zugunsten des höheren Wertes für die Männer erscheint, aber dass der Unterschied bei den Tests mit zweiäugigen Weitteifer, obleich in derselben Richtung, nur sehr gering ist.

GEORGE

STUDIES OF MOTOR FUNCTION IN SCHIZOPHRENIA: I. SPEED OF TAPPING*

From the Psychology Department of the Worcester State Hospital and the Memorial Foundation for Neuro-Endocrine Research

DAVID SHAKOW AND PAUL E. HUSTON

INTRODUCTION

As part of an extensive investigation of schizophrenia from different points of view (6), a study of various motor functions was undertaken. The series of experiments included studies of motor learning as measured by both the pursuitmeter and the prodmeter, simple and discrimination reaction time, steadiness, and tapping.

The content of the psychological schedule given to the groups under consideration need not be discussed here. It is sufficient to say that a survey of psychological functions at different levels of complexity was made as a preliminary approach to more intensive studies of certain aspects of the psychological nature of schizophrenia. Various motor functions were studied from the comparatively simple patellar tendon reflex latent time to learning activities, such as those involved in the pursuitmeter and prodmeter.

As far as the part of the schedule here under discussion is concerned, there is evidence, both of a clinical and experimental nature, that motor speed is a function which is to some extent disturbed in schizophrenic subjects. At the period of the study the primary concern was the description of a schizophrenic population. The present experiment seemed especially indicated since it involved simple conditions, lent itself to quite accurate measurement, and was concerned with a function which showed relatively little change with practice. It also had a systematic place in the battery we were using.

POPULATION

Table 1*a* describes the various groups which have been used in this study in terms of the constants of the distributions of age,

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, May 17, 1935.

TABLE 1*a*

CONSTANTS PERTAINING TO THE DISTRIBUTION OF AGE, DURATION OF PSYCHOSIS, EDUCATIONAL LEVEL, AND OCCUPATIONAL STATUS OF THE VARIOUS GROUPS USED IN THIS STUDY

	Name	Group code letter	*N*	Age (yrs.) Range	Age (yrs.) Mean & σ_m	Age (yrs.) S.D.	Duration (yrs.)** Range	Duration (yrs.)** Mean & σ_m	Duration (yrs.)** S.D.	Educ. level (grades) Range	Educ. level (grades) Mean	Educ. level (grades) S.D.	Occ. stat. (Barr) Range	Occ. stat. (Barr) Mean	Occ. stat. (Barr) S.D.
Schiz. research	One-period	*R*1	82	16-47	30.8±0.8*	7.0	.2-26	5.9±0.6	5.2	4-16	9.6	2.6	1.5-15.8	7.4	3.5
	Three-period	*R*3	38	19-46	32.8±1.2	7.2	.2-21	6.4±0.9	5.5		9.0	1.3		7.6	3.0
	Four-period	*R*4	12	28-46	35.3		.2-21	6.6		4-14	7.8		3.6-13.3	7.5	
Schiz. non-research	One-period	*NR*1	41	21-50	39.3±1.3	8.1	.3-29	13.1±1.0	6.6	2-12	8.5	3.0	3.0-13.0	7.2	3.4
Schiz. merged	One-period	*M*1	123	16-50	33.6±0.8	8.4	.1-29	8.3±0.6	6.7	2-16	9.1	2.8	1.5-15.8	7.4	3.5
Manic-depressive	One-period	*MD*1	13	23-50	41.1		.1-13.6	3.5		0-12	7.4		3.6-11.3	6.7	
Normals	One-period	*N*1	60	17-49	28.3	6.1				4-20	14.9	5.1	3.7-17.0	12.2	4.5
	Two-period	*N*2	21		28.5	5.7					13.9	6.1		11.9	4.5

*All errors are standard errors. This holds throughout the paper.
**Time elapsed since first hospitalization.

TABLE 1*b*

PERCENTAGE OF INCIDENCE OF SCHIZOPHRENIC DIAGNOSTIC TYPES IN THE VARIOUS SAMPLES

	Name	Group code letter	*N*	Par.	Heb.	Ind.	Cat.	Unc.	Mix.	Sim.
Research	One-period	*R*1	82	17	28	12	15	18	5	5
	Three-period	*R*3	38	24	24	18	13	13	3	5
	Four-period	*R*4	12	17	8	25	0	25	8	17
Non-research	One-period	*NR*1	41	24	17	37	0	22	0	0
Merged	One-period	*M*1	123	19	24	20	10	19	3	3

duration of hospital stay, educational level,[1] and occupational status as measured by the Barr Scale (2, pp. 67-69).[2] The percentage distribution of the diagnostic types is given in Table 1*b*.

There are eight groups to be considered in this study, all male and white. Five are schizophrenic, one manic-depressive, and two are made up of normal subjects. Among the schizophrenic groups there are three which are known as "research" groups since the patients included belonged among those under investigation on the Research Service of the Worcester State Hospital. Of these, there is one group made up of 82 subjects (*R*1) who had one period of examination, another group of 38 (*R*3) from among the 82 who had three periods of examination with three-month intervals between sessions, and a third group of 12 (*R*4) from among the 38 who had four periods of examination, with three-month intervals between the first and second, and second and third periods, and approximately four months between the last two. A fourth schizophrenic group is not drawn from the Research Service but from other parts of the male population of the hospital. This is known as the "non-research" group (*NR*1) and consists of 41 subjects who had only one period of examination. A fifth schizophrenic group is called the "merged" group (*M*1) and consists of the 82 research (*R*1) and 41 non-research patients (*NR*1). The manic-depressive group (*MD*1) consists of only 13 subjects. There are two normal groups, one (*N*1), consisting of 60 subjects and having a single period of examination, and another (*N*2) from this group consisting of 21 subjects who had two periods. The interval between these two sessions was approximately four months. (For purposes of special analysis a group of 26 schizophrenic patients and 26 normal subjects matched for age, educational level, and Barr rating were compared.)

The method of selection of the research group (*R*1, *R*3, *R*4) is discussed elsewhere (6, p. 392). The patients in the non-research group were added to make the age distribution more typical of the schizophrenic population of the hospitals in Massachusetts. The

[1]School grade completed is used. In the case of post-grammar school attendance at trade schools, business schools, etc., two years of such attendance is considered the equivalent of one year's attendance at a regular high school.

[2]This scale, although quite unsatisfactory, is probably the best means available for rating occupations.

manic-depressive group is a rather heterogeneous one consisting at the time of the experiment of five patients who might be considered improved: three from a depressed state, one from a manic state, and one from a mixed state; three who were mildly depressed; two who were in a deeper depression; and three who showed a mixed picture. Thus, the subjects were either in an improved or a depressed state.

The main points to be brought out from Table 1*a* are that the normal subjects are younger and of considerably higher educational and occupational status than the patients. Our group of schizophrenics average younger and the psychosis is probably of lesser duration[3] than is true of schizophrenic patients in general in Massachusetts. It must be remembered that only patients below 50 years were used in our study. This would in part account for the differences.

The distribution of the diagnostic types of schizophrenics is given in Table 1*b*. It is impossible to determine the adequacy of the sampling since considerable disagreement, due to varying standards, is to be found in the literature. The diagnoses were made first by the hospital psychiatrists and then by the psychiatric staff of the Research Service, most of the patients being diagnosed repeatedly. It has been found necessary to place a considerable proportion of the patients in other than the four classic types. The latter and the unclassified and mixed designations carry the usual implications. In addition it has seemed desirable to set up another type called the "indeterminate" which is applied to patients who have been definitely schizophrenic at one time but who have apparently lost their type symptoms. They are usually patients who are in some terminal static condition, frequently deteriorated, and usually quite "institutionalized."

Apparatus and Procedure

Among mechanical devices for studying tapping there are two main kinds—the telegraph key and the tapping-plate-stylus arrange-

[3]Duration of psychosis is for our present purposes considered to be the period elapsed since initial induction into a hospital for mental diseases, whether all the time was spent in such an institution or not. Although this is certainly not a satisfactory measure, it is the most satisfactory and objective we have found after experimenting with numerous others.

ment. It is not our purpose here to enter into a detailed criticism of these devices.[4] It is sufficient to point out that the major difficulty of the telegraph key is that it lends itself to "trilling," thus allowing an advantage to one who becomes acquainted with this possibility. The major disadvantage of the tapping plate is that it allows for considerable variety of movement, since the amplitude of rise and the extent of movement of the stylus from the block is not controlled. Other disadvantages are the possibility of "scraping" (side-to-side) movements and still some possibility of trilling. In order to eliminate these disadvantages of the plate Dunlap (2) devised a two-plate tapping board which requires alternate tapping on each of the plates. Some experimentation with this apparatus left us rather dissatisfied since it seemed more cumbersome than necessary and also because some difficulty was experienced with sparking.

In order to avoid these difficulties an apparatus was devised which combined the advantages of the telegraph key with those of the two-plate tapping board and largely eliminated its disadvantages. The major disadvantage not eliminated was the variation in the degree of amplitude of lift of finger from the keys. (It is possible that this is a greater problem abstractly than it is practically, for it was noted that subjects very soon tended to use minimal excursions.)

The tapping block consists of two telegraph keys side by side, two inches apart at the levers. They are mounted on a board ½ inch high and entirely enclosed except for the ends of the key levers. The round buttons of the conventional keys were replaced by square ones—1⅜ inches on each side, thus giving a greater tapping area. The nearest distance between these buttons is ⅝ inch. A wooden partition ½ inch wide with rounded edges projects 5/16 inch above the level of the key buttons. The distance of the inner edge of the key buttons from the partition is 1/16 inch. The partition makes it necessary for the subject to lift the finger slightly in moving from one key to another, and makes impossible the development of any sliding motion from key to key. The excursion of the key

[4]Some discussion of the relative merits of the two types may be found in Ream (8).

is between 1/32 inch and 1/16 inch. The key springs were adjusted to give minimal tension and yet sufficient for the current to be broken as soon as pressure was released.

There are also two major ways of recording taps. It may be graphic or some form of counter may be used. The former is the more satisfactory since it gives a detailed record of each tap. It is the only way of measuring the regularity of tapping during the tapping period. The counter method gives merely the number of taps during the set period. Since a long series of experiments was contemplated, it was considered impracticable to use the more elaborate graphic method. A counter was used which could record 10 taps a second, which brought it within the range of maximum tapping rate.

In order to eliminate the inaccuracies of the stop watch[5] in timing the tapping period, a device consisting of rotary switches and brushes was used. The circuit may be seen in Figure 1.

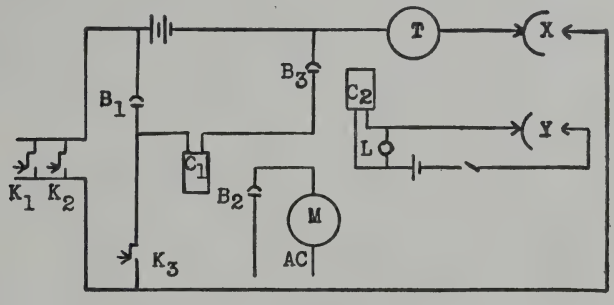


FIGURE 1
TAPPING CIRCUIT

The operation of the tapping circuit is as follows: The experimenter holds down his key (K_3) until the subject has started to tap, and then K_3 is immediately released. This permits current to flow through the circuit-breaker (B_3) which is closed and through the solenoid coil (C_1). The action of C_1 is to close circuit-breakers B_1 and B_2. With B_1 closed, a circuit is maintained through C_1. B_2 when closed starts the Telechron synchronous motor (M) which

[5]Cf. Ream (8, pp. 298-299). He found an average error of .212 sec. (MV = .067) on the part of five observers in estimating a five-second interval.

drives the rotary switches (X and Y). After ¾ of a second the circuit becomes completed through X and the counter (T). The taps made by the subject are now recorded. The rotary contacts and brushes on X are so arranged that taps are counted for five seconds. About ½ second after this interval, X is broken and Y makes contact, momentarily completing another circuit through solenoid coil C_2. This coil operates the circuit breaker B_3 which opens the circuit through C_1. By this means B_1 and B_2 are both opened, the latter stopping the motor. A lamp (L) indicates this event to the experimenter. The whole cycle repeats itself by again closing K_3 and tapping on K_1 or K_2. It will be seen that this scheme permits a constant "warming-up" interval of ¾ of a second, and records taps for a constant period of five seconds without any timing error, unless the experimenter holds K_3 down more than ¾ of a second. If desired K_3 may be introduced in the counter circuit. It would then be necessary for the experimenter to keep it closed during the 5¾ seconds. K_3 was retained in its present position, however, since the method of operation was thought to be somewhat simpler for the experimenter. The arrangement likewise does not require the subject to start to tap at a given signal. The cycle is initiated by the subject.

It will be noted from the above that the interval selected was five seconds. Since we were primarily interested in a test of speed of movement, it was necessary to select a period which gives the maximum number of taps. As nearly as could be determined from the studies of Wells (13), and Benedict, Miles, Roth, and Smith (1, pp. 186, 622), a five-second interval gives the optimum results. After five seconds, fatigue or other factors apparently enter to a considerable extent. Ream (8, p. 303) showed that within a five-second interval there is little change from second to second (1: 7.2; 2: 7.4; 3: 7.2; 4: 7.1; 5: 7.1—18 observers).

The problem of when within the interval to start the recording of the tapping arises. We have followed Ream's method and have not counted the first ¾ second of tapping. Due to the inertia of getting started and the irregularity of the beginning of the tapping period, this part does not give a good sample.[6] The preliminary period was kept constant automatically.

[6]Other experiments done with schizophrenic subjects indicate the importance of analyzing the "practice" period, however.

Ten trials, with two preliminary practice periods, were selected as the optimum number. It was felt that this was most likely to meet the exigencies of the situation when the battery of tests to be administered, the nature of the population involved, etc., were taken into consideration. Ream seems to think that 20 five-second trials is the optimum. However, we were limited in time and since his data do not show any substantial increase after ten trials this number was adopted. He used a "few moments" rest between trials. Our rest periods were 15 seconds in length.

The directions were essentially as follows: The subject was seated in front of the tapping block with the forearm of his preferred hand resting on the table, and made as comfortable as possible. He was then told:

> This is a simple device which is intended to find out how quickly you can tap back and forth with the forefinger on the two keys. [Demonstrated] Let your forearm rest on the table in a comfortable position, grasp the edge of the tapping instrument with the other hand, and lean slightly forward. [Demonstrated] Make the tapping movement mostly from the wrist rather than from the elbow. [Demonstrated] Remember to tap as rapidly as possible. At the signal "Ready!" place your arm in a position to tap and at the signal "Go!" start tapping and continue until the signal "Stop!" The more completely you relax between trials the faster you can tap. Ready—Go!

Contact is made with Switch K_3 until the subject has begun to tap, and then it is released. When the light (L) came on, indicating that the counter (C_2) had discontinued recording, the subject was instructed to stop. Two practice trials were given during which the subject was observed carefully to make sure that instructions were followed, that he was using the most economical movements, and that any tension which existed was as much as possible eliminated. The experiment was performed in a well-lighted room. The subject was encouraged before and between trials as much as possible to do his optimum work. The experimental sessions were held mornings (9:00-11:00 A.M.) and afternoons (1:30-4:00 P.M.)—approximately half of each of the groups used coming at each time. In the portion of the population with whom the experiment

was repeated one or more times, the additional sessions were always held at approximately the same time of the day as the initial session.

Results

1. *Introduction.* Because of the variety of groups and the number of ways in which the data are treated, the best way of presenting these is first to give the findings for the periods taken as units, and then those for the separate trials. The discussion of the reliability of the findings and the factors which might account for any existing differences can then follow together with the discussion of the results. It will facilitate the assimilation of the data if the nature of the groups as described in Tables 1*a* and 1*b* and the following definitions are kept in mind:

> Tap—Single tap on telegraph key;
> Trial—Five seconds consecutive alternate tapping on tapping block;
> Session or period—Ten trials of five seconds each, with two preceding practice trials of five seconds each.

2. *Periods.* Table 2 gives the means of the mean number of taps for the ten trials and the means of the best of the ten trials together with various other measures, mainly of variation, on 82 research schizophrenics, 43 non-research schizophrenics, both groups merged (125), 13 manic-depressive patients, and 60 normal controls. There are also data on the mean performance of 26 schizophrenic patients and 26 normals matched for age, education, and Barr rating.

The difference of 9.4 taps (19.5 vs. 28.9)[7] between the means of means of the schizophrenic patients—the merged group is used throughout unless otherwise specified—and the normals is significant (t[8] = 11.09, $P = < .01$).[9] The mean of the manic-depressive

[7]Cf. Johnson, B. (7). A computation from her data for three male subjects gives a similar mean. She used a double plate tapping board. With psychotic subjects it might be profitable to use some such technique of tension recording as that of Johnson. Comparisons with other data on normals in the literature are difficult because of differences in apparatus and technique.

[8]Because of the one small sample (the manic-depressive) compared, it was necessary to use Fisher's "t" (4, pp. 111-151) as the criterion for the significance of the differences between means. The use of the $D/\sigma_{diff.}$ is not legitimate for such small samples.

[9]We have been able to discover three previous studies involving schizo-

TABLE 2

CONSTANTS OF THE DISTRIBUTIONS OF MEAN NUMBER OF TAPS FOR TEN TRIALS, RANGE OF TEN TRIALS, BEST SCORE OF TEN TRIALS, POOREST SCORE OF TEN TRIALS FOR VARIOUS GROUPS FOR SESSION I

Variable	Group	N	Min.	Max.	Range	Mean & σ_m	S.D. & $\sigma_{S.D.}$	C.V.
Mean of 10 trials	Research I	82	7.5	34.5	27.0	18.8± .64	5.76±.45	30.6%
	Non-research	43	9.6	29.6	20.0	20.0± .76	5.01±.54	24.2%
	Merged	125	7.5	34.5	27.0	19.5± .50	5.58±.35	28.6%
	Matched schiz.	26	7.5	33.6	26.1	19.6±1.27	6.34±.89	
	Manic-dep.	13				24.8		
	Normals	60	21.5	39.2	17.7	28.9± .55	4.24±.39	15.4%
	Matched normals	26	21.5	35.0	13.5	27.7± .76	3.78±.53	
Range of 10 trials	Research I	82	2.0	17.0	15.0	7.8± .35	3.20±.25	40.9%
	Non-research	43	2.0	14.0	12.0	5.6± .34	2.20±.24	39.3%
	Merged	125	2.0	17.0	15.0	7.1± .28	3.08±.19	43.6%
	Manic-dep.	13				7.4		
	Normals	60	2.0	10.0	8.0	5.2± .23	1.77±.16	33.6%
Best of 10 trials	Research I	82				22.7± .63	5.74±.45	25.3%
	Non-research	43				23.5± .72	4.69±.51	20.0%
	Merged	125				23.0± .48	5.41±.34	23.6%
	Normals	60				31.5± .60	4.61±.42	14.6%
Poorest of 10 trials	Normals I	59				26.1± .54	4.14	

group falls between the above two (24.8),[10] and although the group is small ($n = 13$) the differences are also significant each way when tested by the same measure. It is 2.98 with a P of $<.01$ between normals and manics, and 3.12 with a P of $<.01$ between schizophrenics and manics.

A comparison of the means of the best of ten trials gives a difference of 8.5 taps (23.0 vs. 31.5) between schizophrenic and normal subjects, which is also significant. (The significance of the differences has not been computed since it is practically the same as that of the means of means.) It might be pointed out that the schizophrenic patients do not reach the level of the mean performance of the normals on their best trial (23.0 vs. 28.9) and, what is even more striking, do not reach the level of the normals' average poorest performance on the mean of their best trials (23.0 vs. 26.1). The $D/\sigma_{diff.}$ is 4.31, which is significant.

The consideration of the various measures of dispersion (Table 2) is now indicated. The standard deviations[11] show a distinct tendency to be greater in the patients than in the normals. This is corroborated both by the ranges of the means in the groups and the means of the individual ranges. The latter measure indicates a great individual variation as well. The standard deviation of the schizophrenic group is 5.58, whereas that of the normals is 4.24. The significance of the difference of the standard deviations is, however, only on the border of significance ($D/\sigma_{diff.} = 2.6$). The

---

phrenic subjects. The first two are those of Boring. (1*a*) and Gatewood (5) who used a pencil-paper procedure with a single 30-second tapping period. Because of the difference in technique comparisons are difficult to make. The other is that of Wells (12) on manic-depressives. A single tapping key was used so that the comparisons which can be made are limited. In the investigation six dementia-praecox subjects and ten normal subjects were also examined. From the graph he presents on page 47 it is possible to compute roughly the mean number of taps per second for the first five-second interval for the dementia-praecox subjects and the normals. These are 6.4 and 7.0 respectively. The figures for our groups are 3.9 and 5.7 respectively—a much greater relative difference than is found in Wells' groups. Whether this is due to the greater complexity of our procedure or to selective factors in the groups cannot of course be determined.

[10]No relationship between psychiatric condition and score could be determined by inspection of the data on these patients.

[11]Although the coefficients of variation are presented in the tables, they will not be discussed because of the undetermined meaning of this statistic.

difference between the means of individual ranges is definitely significant ($D/\sigma_{diff.} = 6.2$). The consistency of the tendency is also seen in the standard deviations of the individual ranges, which are 3.08 and 1.77 for the patients and normals respectively. The difference is significant ($D/\sigma_{diff.} = 4.7$). A consideration of the standard deviations of the best trials shows the expected lesser difference between the two groups, namely, only .80. It might be pointed out that, although the mean performance of the manic-depressives is more nearly like that of the normals, the mean of the individual ranges is distinctly higher and even above that of the schizophrenics (7.4 vs. 7.1, 5.2 being the normal). Because of the small number of subjects, this is the only measure of variation we have computed in the manic-depressive group.

Thus far our results show a significantly poorer mean performance of schizophrenic subjects together with a consistently greater group and individual variation than found in normals. The manic-depressive group, within the limits of generalization permitted with a rather small number of subjects, seems to fall in between the other two groups in mean performance.

The material available on the repeated sessions might now be considered. When the groups are compared by periods (Table 3), it is found that there are 38 schizophrenics who have had three periods, 12 of these also having a fourth, and 21 normals who have had a second period. The 60 normal controls will also be used for comparison here, together with some data which are available on the first period of the 21 normal subjects having a second period. Since the manic-depressive patients were available for only one session, no report can be made on this group.

A comparison of the means of the means for each of the three periods shows scores of: I—19.3; II—20.6; III—22.3. The score of the fourth period for the 12 subjects is also 22.3. Except for the fourth period, a consistent tendency upwards is shown. The significance of the difference between the mean scores was computed by Student's Z method.[12] The Z's are .567 and .577 for I and II,

[12]The usual formula $\sqrt{\sigma m_1^2+\sigma m_2^2-2r_{12}\sigma m_1\sigma m_2}$ was not used. Because of the wide fluctuation of an r based on only 38 cases, it was deemed advisable to use a method which discounts the correlation present—in this case Student's Z (10).

TABLE 3

CONSTANTS OF THE DISTRIBUTIONS OF MEAN NUMBER OF TAPS FOR TEN TRIALS, RANGE OF TEN TRIALS, BEST SCORE OF TEN TRIALS, RANGE OF MEANS OVER THREE SESSIONS FOR VARIOUS GROUPS FOR THREE AND FOUR SESSIONS

Variable	Group	N	Min.	Max.	Range	Mean & σ_m	S.D. & $\sigma_{S.D.}$	C.V.
Mean of 10 trials	Research I	38	10.0	33.6	23.6	19.3±.81	4.98±.57	25.8%
	Research II	38	13.1	34.8	21.7	20.6±.72	4.44±.51	21.6%
	Research III	38	15.2	36.5	21.3	22.3±.67	4.14±.47	18.6%
	Research IV	12				22.3*		
	Normals I	60	21.5	39.2	17.7	28.9±.55	4.24±.39	15.4%
	Normals I*a***	21				28.9		
	Normals II	21	23.4	37.5	14.1	28.8±.72	3.29±.51	11.0%
Range of 10 trials	Research I	38	2.0	14.0	12.0	7.5±.53	3.28±.38	43.7%
	Research II	38	2.0	15.0	13.0	6.5±.46	2.84±.33	43.7%
	Research III	38	2.0	12.0	10.0	6.7±.39	2.38±.27	28.2%
	Research IV	12				5.9		
	Normals I	60	2.0	10.0	8.0	5.2±.23	1.77±.16	33.6%
	Normals II	21	2.0	11.0	9.0	5.3±.46	2.10±.32	40.0%
Range of means over 3 sessions	Research	31	0.1	12.7	12.6	3.94±.43	2.64±.30	67.0%
Best of 10 trials	Research I	38				23.2±.83	5.10±.59	22.0%
	Research II	38				24.1±.62	3.85±.44	16.0%
	Research III	38				25.5±.69	4.24±.49	16.6%
	Normals I	60				31.5±.60	4.61±.42	14.6%

*The mean of the third period for the 12 who also had a fourth period was 22.3.

**Normals I*a* is the group of normals who had a second period. It is made up of the same individuals as those in Normals II.

and II and III respectively, to which P's of less than .01 correspond. There is no difference between periods III and IV. Due to the fact that the experiment was not carried beyond this point, it is impossible to say whether or not a level really was reached.

The normals show no change in the second period over the first—the respective scores being 28.8 and 28.9 for the group of 21 normals. (The total group of 60 had shown a score of 28.9 in the first period—a gratifying consistency which points to the good sampling in the smaller group.) Unfortunately, it was not possible to obtain three-period records on the normals. It will be noted that the schizophrenic patients' score in Period III is still significantly below the normals' score for Period I (22.3 vs. 28.9), the $D/\sigma_{diff.}$ being 7.6. This is, of course, also true of the score for Period IV of the patients.

A similar comparison of the mean of the best trials shows some increase of the score from period to period in the patients, but somewhat less than takes place in mean score. It is interesting to note that even the mean score of best trials of Period III for the schizophrenics does not reach the level of the mean of the mean scores for normals for the first period (25.5 vs. 28.9). The difference is still significant ($D/\sigma_{diff.} = 4.0$). If we compare patients' best scores for Period III with the normals' poorest for Period I, we find that the difference is no longer significant (25.5 vs. 26.1), but still slightly lower.

The measures of variation for the different periods offer material for consideration. The range of the means for the schizophrenics decreases consistently from period to period, but is still higher in Period III than that of the normals for Period I (21.3 vs. 17.7). (Period IV cannot legitimately be compared since the population is much smaller.) The standard deviations show the same tendency to decrease, and when the third period of the patients is compared with the first period of the normals ($n = 60$), little difference is found. The mean of individual ranges reveals a tendency to decrease from Period I to Period II, but remains constant in Period III, in all cases being higher than that for the normals of the first period. The standard deviations of the individual ranges show a consistent drop, but never reach the low level of the large group

of normals. The normals show the same tendency for the standard deviations to drop from Period I to Period II.

A comparison of the variation in the best trials shows a greater standard deviation in Period I as compared with both Periods II and III. The differences are not, however, significant.

The results obtained in the different diagnostic types of schizophrenia (Table 4) may now be indicated. In the group of 123

TABLE 4

CONSTANTS OF THE DISTRIBUTIONS OF MEAN NUMBER OF TAPS FOR TEN TRIALS OF DIAGNOSTIC TYPES OF SCHIZOPHRENIA FOR SESSION I

Group	N	Mean of Means & σ_m	S.D. & $\sigma_{S.D.}$	C.V.
Paranoid	24	24.4±0.94	4.60±.66	18.8%
Hebephrenic	30	19.5±0.94	5.14±.66	26.4%
Catatonic	12	13.5±1.06	3.68±.75	27.3%
Indeterminate	25	19.1±0.61	3.06±.43	16.0%
Unclassified	24	20.3±1.26	6.16±.81	30.4%
Simple	4	19.0		
Mixed	4	18.1		
Total	123	19.9±0.50	5.58±.35	28.6%

patients who had one period, it is seen that the performance of the catatonic group is distinctly the poorest and that of the paranoid group is distinctly the best. The difference is significant—in fact the distributions of these two groups show practically no overlapping and would indicate that one is dealing with two different populations. The paranoid group, however, still shows a significantly lower score than the normal group ($D/\sigma_{diff.} = 4.0$). The indeterminate, unclassified, and hebephrenic groups fall in between at about the same level. Frequency distributions of these three types show a marked degree of superimposition. (The mixed and simple types with four cases each also fall at about the same level.) The $D/\sigma_{diff.}$ between the paranoid and unclassified types is 2.6—on the border of significance. The difference between the paranoids and the hebephrenics is significant ($D/\sigma_{diff.} = 3.7$), as is that between the paranoid and the indeterminate type ($D/\sigma_{diff.} = 4.7$).

A consideration of the variation in these groups shows the least variation, as measured by the standard deviation, for the indeterminate and catatonic groups, and the most for the unclassified and

hebephrenic groups. The difference in standard deviation between the extreme groups (indeterminate 3.06 vs. unclassified 6.16) is significant ($D/\sigma_{diff.} = 3.4$).

Taking the types through the entire experiment, which covered three periods, necessarily limits us to the small group involved ($n = 38$) and in such an analysis we cannot do more than follow tendencies. It would be more satisfactory, however, if this particular analysis were made during the discussion of the performance on separate trials, so it may be postponed until then.

As pointed out earlier, in the discussion of the population, there were two groups of schizophrenic patients studied. One was the research group who were on special study, consisting of 82 subjects for the first period, and the other was a group of 43 patients taken from the hospital at large and not on special study. Since the latter was primarily chosen to round out the former group, there are distinct differences in the constitution of the group, as has already been pointed out. Despite these differences, the constants (Table 2) for this group are quite similar to those of the research group. Nowhere are the differences significant. In general the variation seems to be less than that found in the research group.

3. *Trials.* Since the results obtained on successive trials show some differences, they are presented in Table 5. There is no overlapping in any of the trials on the part of the schizophrenic groups (research and non-research), manic-depressives, and normals. The standard deviations of the normal group and the schizophrenic merged group show the same absence of overlap. (The variation of the manics is not reported upon because of the small number of cases.) The consistency of performance in all three groups over the period is striking. A consideration of the differences between the means of the normals and that of the schizophrenic merged group reveals that they are significant for each trial.

When an examination is made of the trial in which the best score[13] is most frequently made, it is found that in the schizophrenics 16 per cent attain it on the second trial, in the manic-depressives 23 per cent on the ninth trial, and in the normals 27 per cent on the

[13]Since the range is small, maximum values frequently recur once or twice in the course of ten trials. The *first* trial on which the maximum score occurs is the one used here.

TABLE 5

MEANS AND STANDARD DEVIATIONS BY TRIALS

Single Session Groups

Trial No.	Research (82) Mean	S.D.	Non-research (41) Mean	S.D.	Merged (123) Mean	S.D.	M.D. (13) Mean	Normals (60) Mean	S.D.
1	18.8±.66	6.0±.46	20.3±.75	5.0±.53	19.3±.51	5.7±.36	24.3	28.5±.52	4.1±.37
2	19.5±.67	6.1±.48	21.4±.47	4.4±.47	20.2±.51	5.7±.36	24.2	29.1±.55	4.3±.39
3	19.1±.71	6.4±.50	21.1±.76	5.0±.54	19.8±.54	6.1±.38	24.2	28.9±.59	4.6±.42
4	19.5±.72	6.6±.51	21.3±.71	4.7±.50	20.1±.54	6.1±.38	24.9	28.9±.63	4.9±.45
5	19.3±.66	6.0±.46	21.1±.83	5.5±.59	20.0±.52	5.8±.37	25.4	28.8±.52	4.0±.37
6	19.3±.75	6.8±.53	21.1±.89	5.9±.63	19.6±.56	6.3±.40	25.0	28.8±.63	4.9±.44
7	19.3±.61	5.5±.93	21.9±.83	5.5±.59	20.2±.51	5.7±.36	25.8	28.9±.57	4.5±.41
8	19.5±.68	6.2±.48	20.8±.79	5.2±.56	20.0±.53	5.9±.37	24.1	28.8±.64	5.0±.45
9	19.4±.67	6.1±.48	20.8±.90	5.9±.64	19.8±.54	6.1±.38	25.2	28.3±.57	4.4±.40
10	19.7±.68	6.2±.48	21.1±.85	5.6±.60	20.2±.54	6.0±.38	24.5	28.5±.60	4.7±.42

TABLE 6

MEANS AND STANDARD DEVIATIONS BY TRIALS

Repeated Session Groups

Trial No.	Research I (38) Mean	Research I (38) S.D.	Research II (38) Mean	Research II (38) S.D.	Research III (38) Mean	Research III (38) S.D.	Research III*a* (12) Mean	Research IV (12) Mean	Normals I*a* (21) Mean	Normals II (21) Mean
1	19.8± .85	5.3±.60	20.6±.70	4.3±.49	21.3±.73	4.5±.52	20.0	21.3	27.5	29.8
2	20.2± .92	5.7±.65	21.6±.74	4.6±.52	22.3±.75	4.7±.53	21.3	22.3	28.0	29.8
3	19.5±1.04	6.4±.73	20.4±.75	4.6±.53	22.8±.70	4.3±.49	22.4	21.8	28.1	29.3
4	20.0±1.04	6.4±.73	21.5±.72	4.5±.51	22.8±.81	5.0±.57	22.1	23.2	28.3	28.8
5	19.5± .81	5.0±.57	20.9±.78	4.8±.55	23.1±.63	3.9±.45	23.3	21.5	27.7	28.4
6	19.9± .99	6.1±.70	21.1±.96	5.9±.68	22.5±.75	4.6±.53	23.0	22.9	28.2	28.1
7	20.2± .76	4.7±.54	21.5±.70	4.3±.49	23.1±.69	4.2±.49	22.6	23.1	28.2	28.3
8	20.3± .83	5.1±.59	21.6±.80	5.0±.57	22.9±.68	4.2±.48	22.8	22.4	28.0	28.3
9	20.0± .75	4.6±.53	21.5±.65	4.0±.46	22.9±.64	3.9±.45	22.7	22.2	27.7	28.9
10	19.9± .85	5.3±.60	21.7±.73	4.5±.52	23.2±.71	4.4±.50	22.4	22.8	27.9	28.4

second trial. In the last group 73 per cent of the subjects have attained their highest score in the first five trials, in the schizophrenics 58 per cent, and in the manic-depressives 54 per cent.

An analysis of the occurrence of reversals of any trial over the preceding trial shows little difference between schizophrenic patients and normals. Because of the nature of the function—one which involves little learning—the problem of reversals does not seem particularly significant.

An analysis of the data relative to achievement in the repeated session groups by trials shows the following (Table 6): In the schizophrenics, the third-period score for each trial shows a slight but consistent increase (without exception) over that of the second period and the second the same over the first period. The fourth period shows variations each way, but almost identical scores. Each period shows a marked consistency of performance within itself. The curve followed through shows a slight upward rise which levels off at Period III.

A consideration of the variation from trial to trial (Tables 5 and 6) in the schizophrenics shows a tendency for the standard deviations to remain fairly constant within periods, but to decrease from period to period. The level of the standard deviations for the normals runs at about that of the third period of the patients.

An analysis of the type scores by trials for the first period is presented in Table 7. From an examination of the table it will be seen that what was true for the mean of ten trials is true for the means of each of the ten trials—the paranoids are the best, the catatonics the poorest, and the unclassified, hebephrenic, and indeterminate groups fall in between. There is no overlapping except in the case of the middle three types.

An analysis of the first appearance of best trials shows that the catatonics seem to distinguish themselves from the others in that 67 per cent of the cases obtain the best scores during the first two trials. The next highest group is the indeterminate, with only 28 per cent.

A number of interesting points come out from a study of the types in the three-period group (Figure 2). Because of the small number of cases in each diagnostic type the results must not be taken to represent more than trends, but because of the consis-

TABLE 7

MEAN NUMBER OF TAPS BY TRIALS
Merged Group (By Diagnostic Type)

Trial No.	Heb. ($n = 30$)	Cat. ($n = 12$)	Par. ($n = 24$)	Indet. ($n = 22$)	Uncl. ($n = 20$)	Simp. ($n = 4$)	Mix. ($n = 4$)
1	18.3	14.0	24.1	18.6	20.0	18.5	16.5
2	19.3	13.6	24.0	18.8	21.0	18.0	16.3
3	18.9	12.8	24.3	19.5	20.3	18.8	15.3
4	20.1	12.3	24.7	18.8	20.4	19.8	16.5
5	20.1	13.8	24.9	18.9	21.0	17.3	16.5
6	19.8	11.3	24.5	18.4	21.4	19.5	18.3
7	19.4	13.9	24.5	18.7	20.7	19.0	18.3
8	20.1	14.4	24.7	18.6	21.0	19.3	17.5
9	21.5	14.3	24.5	19.1	20.9	20.0	15.8
10	20.1	14.0	24.3	19.1	21.0	18.8	18.8

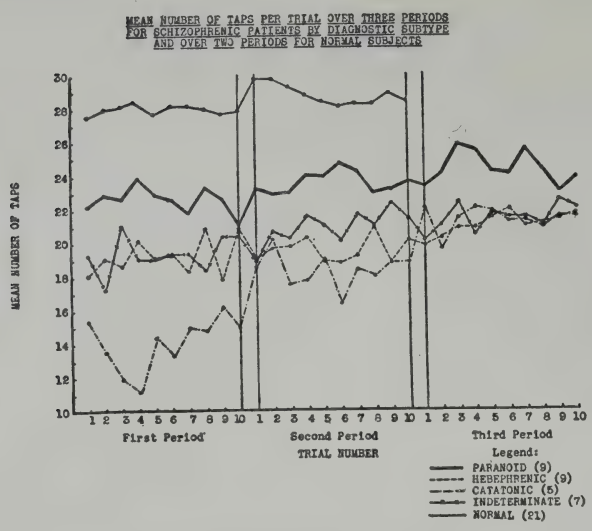


FIGURE 2

tency of these results with those for the single period there seems some ground for considering them valid. An examination of Figure 2 shows: (1) A difference between types which tends to disappear with repetition, i.e., the curves tend to come together in the third period. (2) The most marked improvement is shown by the catatonics. (It might be mentioned in passing that this improvement does not seem correlated with improvement in psychiatric condition.) (3) The differentiation between the hebephrenic and indeterminate types on one hand, and the catatonics on the other, is wiped out by the third period. The hebephrenic and indeterminate types stay close together throughout and show a slight rise on the whole. (4) The paranoids show a consistent performance throughout the three periods, always maintaining their superiority and in no trial being overlapped by the other groups.

4. *Consistency.* Before proceeding to a discussion of the factors which might account for the differences we have found among the various groups, it is necessary to say something about the consistency of the function involved and the consistency of the present technique in measuring this function. We are concerned with the problem

of the "reliability" of the tapping method used by us. This can be determined both by correlation measures which take into account the constancy of the relationship (without the regression equation one cannot tell whether the levels have been maintained), and by absolute measures of differences between scores of various kinds which are roughly indicated by the means. As far as self-correlation measures go, it seems correct to generalize as follows: Low relationship may mean either that we are dealing with a consistent function and an inconsistent measuring device, an inconsistent function and a consistent measuring device, or an inconsistent function and an inconsistent measuring device. High relationship could mean only one thing—that both the function and the measuring device are consistent. The other possibility, that of an inconsistent function with a compensating measuring device, is very improbable. We shall consider the self-correlations with these possibilities in mind.

Taking the whole groups there are various measures *within* sessions—measures of short-period consistency—which enable one to judge the reliability of the technique. The correlation coefficients[14] which are found are given in Table 8.

TABLE 8
SELF-CORRELATIONS
Within Periods

Measures correlated	Schiz. merged (123)	Schiz. III (38) I	II	III	Normals (60)
Mean 1st 5 trials/ Mean last 5 trials	.95±.01				.91±.02
Mean odd trials/ Mean even trials	.94±.01				
Trial 5/ Trial 6	.86±.02				
Trial 1/ Trial 10		.71±.08	.73±.08		

It will be noted that there is a high degree of internal consistency within the sessions when analyzed either by mean of first five trials against that of last five trials, by mean of odd trials against mean of even trials, or when one trial is taken against a successive one (trial 5 vs. trial 6). The coefficients are .95, .94, and .86, respectively. When the first trial is taken against the last in the

[14]All are product moment coefficients.

smaller group of patients, coefficients of .71 and .73 are found for the first two periods respectively. An examination of the regression coefficients ($0.978x$, $0.964x$, $0.920x$) for the first five/last five, odd/even, 5th/6th correlations shows that there is practically a one-to-one relationship.

An analysis of these in the normal subjects shows more or less the same thing. Only the coefficients for the mean of first five/mean of last five trials was computed. This was .91. The others, to judge from inspection of the data, are undoubtedly very similar.

A measure of reliability which is perhaps more satisfactory is the "test-retest" correlation—that between *successive* sessions. Table 9 gives the correlations for the available patient group—that of the

TABLE 9
SELF-CORRELATIONS
Between Periods

Measures correlated	Schiz. III (38)	Normals (20)
Mean I/ Mean II	.88±.04	.74±.09
Mean I/ Mean III	.73±.08	
Mean II/ Mean III	.79±.06	
Range I/ Range II	.44±.13	
Range I/ Range III	.41±.14	
Range II/ Range III	—.07±.16	
Trial I_1/ Trial II_1	.69±.09	
Trial I_1/ Trial II_{10}	.71±.08	
Trial I_1/ Trial III_1	.68±.09	
Trial II_1/ Trial III_1	.69±.09	
Trial I_{10}/ Trial II_1	.74±.08	
Trial I_{10}/ Trial III_1	.59±.11	
Trial II_1/ Trial III_1	.69±.09	
Trial II_{10}/ Trial III_1	.66±.09	
R: II_1/ I_1, 10	.78±.07	
R: III_1/ II_1 10	.73±.08	

repeated schizophrenics. The correlation coefficients for the means are .88, .79, and .73. When it is remembered that there is a three-month interval between the measurements in the first two, and a six-month interval in the last, the coefficients may be considered as fairly high. The only other group on whom data are available is the repeat group of normals, consisting of 21 individuals. The coefficient is .74—lower than that for the patients.

Other measures of test-retest reliability are available, and are

given in the table. The multiple correlation of the first trial of the second period with the first and last trials of the first is .78, and that of the first trial of the third period with the first and last of the second period is .73. The coefficients of the ranges are for individual ranges.

In a comparison of type differences in reliability, the only available data, even approximately satisfactory, are those of the first period. One is therefore necessarily limited to a consideration of the reliability shown *within* a session. An examination of the scatter diagrams for the first session of the merged group shows that there is little to choose between types except that the catatonics seem to have the greatest tendency to diverge from the regression line.

From the data presented on inter-session and intra-session consistency, it would seem that the function is one which remains very consistent through a session and fairly consistent over periods of three and six months.

5. *Sources of Variation.* Having established a number of significant differences between various of the groups and having also established a fairly high degree of reliability of the procedure itself, indicating that differences are probably not due to chance and that there is a fair constancy of the function and its measuring device, a problem which now takes on particular interest is that of the factors which may account for the differences.

One should first want to know whether these differences may possibly be due to factors related to selection which really have little or no relation to the psychosis. Then one should want to consider the possible effect of factors which are related to the immediate experimental situation and, finally, those in which the general psychotic condition seems more particularly involved.

To take the factors outside of the psychotic situation, it might be assumed that a number of these such as age, duration of psychosis, etc., might affect the score. As indicated previously, under the discussion of the population, there were differences in the constitution of the patient and normal groups in certain characteristics. With this in view, let us examine the correlations obtained between the scores and these variables.

What effect does age have on the score? Table 10 gives the correlations. What little relationship there is, is not rectilinear.

TABLE 10

CORRELATIONS OF SCORE AND RANGE WITH CERTAIN SOURCES OF VARIATION

	Group	C.A.	Duration of psychosis	Education	Barr	Cooperation	Adjustment type
Mean Score	*R*1	.35±.10*	.31±.10*	—.08	.03	.66±.06	
	*R*3I	.46±.13*	.35±.14*			.65±.10	.34±.14
	*R*3II	.49±.12*	.34±.15*			.65±.10	.48±.13
	*R*3III	.38±.14*	.52±.12*			.71±.09	.38±.14
	*M*1			—.02	.02	.66±.05	
	*N*1	—.18±.14		.35	.40		
Range	*R*1					—.15±.11	
	*R*3I					—.16±.16	—.25±.15
	*R*3II					—.47±.13	—.52±.12
	*R*3III					.00±.17	—.06±.16
	*N*1	R mean (age, educ., Barr) = .437					

*Correlation ratio.

We find a correlation ratio of .35 in the case of the large research group and the results for the three-period groups are a little higher with correlation ratios of .46, .49, and .38, respectively, for the three periods. The low peak of the curve is between years 30 and 35. The normals show a negligible relationship between age and score ($r = -.18$) although the age range is approximately the same. Even if one were not to consider the small size of the group, the small size of the ratios, and the fact that they *are* ratios, and therefore much less dependable than coefficients, it is reasonable to assume that this relationship is not a direct one. When one examines the scatter diagrams of the patients, it is found that the existence of ratios of any size seems largely determined by the poor scores of the catatonics (who are the youngest group) and a few young hebephrenics, and the medium range of scores of the other groups. We think it is legitimate to say that age is not a factor in accounting for the differences, or at most, a very minor factor. This is corroborated by the fact that taking the diagnostic types separately reveals no relationship with age.

When one approaches the matter of the period of duration of the psychosis (Table 10), one again finds a curvilinear relationship between this factor and score of approximately the same level as that for age. Here the tendency seems to be for no effect to show itself until eight to ten years of duration and then for the scores to decrease with increasing duration. This result for the large group is corroborated by the findings in the three periods of the smaller group. An analysis of the scatter diagram reveals that the minor degree of relationship is determined by the hebephrenics, indeterminates, and paranoids in whom the psychosis is of longer duration.

It will be seen from the above that, although apparently both age and duration of psychosis have a slight effect on score in patients, this is largely an indirect effect since it is influenced by an unavoidable age sampling in types. A consideration of Table 11 will

TABLE 11

MEANS OF AGE AND DURATION OF PSYCHOSIS (IN YEARS) OF DIAGNOSTIC TYPES

	Heb.	Cat.	Par.	Indet.	Uncl.
Age	28.5	26.3	36.0	34.9	30.0
Duration	7.0	3.0	4.5	10.5	4.5

show that there is a marked difference between the types with regard to both age and duration of psychosis.

With respect to education, there seems to be no relation between grade completed and mean score. The coefficients for the two large patient groups ($R1$ and $M1$) are —.08 and —.02, respectively. In the normals, however, the coefficient is higher, .35—still a low degree of relationship.

The Barr ratings give similar coefficients in the three groups, .03, .02, and .40 respectively. For the normal group (where all relations are linear) a multiple correlation of .44 between mean score and age, education and Barr rating, is obtained. It is thus seen that even in this group the relationship is rather low.

But since there is a slight influence of age and duration of psychosis on the patients, and a slight influence of occupational level on normals, and since there are differences in the constitution of the groups—the patients being older, and the normals having higher Barr and educational ratings—it may be that the cumulative effect of these differences accounts for the differences found in means. To check upon this a comparison was made between groups of 26 schizophrenics and normals matched for age, education, and Barr rating. The respective means are 19.6 ± 1.3 and 27.7 ± 0.8, and the respective standard deviations, 6.3 ± 0.9 and 3.8 ± 0.5. It will be seen that the mean of the matched patient group is practically the same as that for the whole group for the first period and that the mean of the matched normal group is only slightly lower (1.2 taps) than that for the large normal group. The $D/\sigma_{diff.}$ here is 5.4, a significant one. The implications drawn from the correlations are thus borne out by the results from matching.

Before proceeding to a discussion of the major factors involved in the immediate experimental situation, it might be mentioned that an analysis of the effects of the time of day of the experiment—morning as opposed to afternoon results—showed no difference in mean or standard deviation in either the schizophrenic or the normal group.

The effect of attitude (mainly as evidenced in cooperation) is probably the most difficult aspect of the problem. There seem to be three levels of performance which are involved here, as in practically all psychological work with psychotic subjects. The frequently

used terms *achievement, ability,* and *capacity* are adopted for these three levels.

By the level of achievement is meant what the subject actually performs in a task—his score. Obviously a psychotic patient's (and to some extent a normal person's) performance in a task which involves motivation and exclusion of response to extraneous stimulation is likely to be below his optimum level. It is affected by his attitude to the immediate experimental situation, his characteristic (long-lasting) attitude, and by the more cryptic and probably more fundamental effects of the psychosis. It is very difficult to distinguish between the latter two, but we mean by the former that which may be recognized by certain symptoms such as apathy, distractibility, psychomotor retardation, etc., and by the latter, the kinds of states which are not discovered except by experimental or thorough clinical investigation.

In practically all of our psychological work with patients—except in the cases where one has either descended to such a simple level that one is dealing with behavior primarily segmental (e.g., patellar tendon reflex) or where one is interested in directly studying the attitudes themselves, one must have some way of determining the quality of the attitudes displayed towards the task in hand. Since an instrumental or experimental way of recording degree of interest, voluntary attention, and distractibility has not as yet been devised, we have been compelled to fall back on observation. To make these observations as objective as possible a rating scale was developed which is used in the case of all of the experiments in which it is relevant. It is a five-point scale which allows for intermediate ratings.

Rating Scale

A Enthusiastic interest in the task; desires to beat own record, to do as well as possible.

B Fairly interested in the task; anxious to do well; very good effort but not enthusiastic.

C Urging necessary—will complete the task, but effort is perfunctory or poor. Doesn't care particularly what the results are—only wishes to get through as soon as possible.

D Only partial results obtained. Patient uncooperative; may

stop in the course of the experiment and refuse to do any more.

E Totally uncooperative; no results.

The value of this rating is that it gives the experimenter's judgment as to the degree to which the *ability* of a subject is reflected in his *achievement*. By ability is meant what the individual is capable of doing at his present level of training and development when both the interfering factors which are peculiar to the immediate situation and those interfering factors, which, if present, are more or less characteristic of his day-to-day condition, are overcome. When, in the opinion of the experimenter, the subject shows neither of these types of behavior in the experimental session, then he is likely to rate the subject *A*, which means that the achievement is in his opinion at an ability level.

It might be pointed out here in passing that the first kind of difficulty—factors of an interfering nature especially connected with the experimental situation—were probably present to only a small degree in this experiment. The evidence we can give on this point is connected with the correlation of the *adjustment type*[15] rating of the psychiatrist and the cooperation rating of the experimenter. The correlations for the three periods (Table 12) are: .41, .50, .32, respectively—fair correlations for the subjective measures with which we are dealing. The standard errors are high and the correlations should be considered with this in view. The emphasis is being placed on the consistency of the tendency rather than on the size of the coefficients. An examination of the scatter diagrams for these correlations reveals that the cooperation ratings were in the great majority of cases better than the adjustment ratings of the same individual.[16] The figures are: for Period I—

[15]A rating scale (3) of the general adjustment level of the patients on the research wards made by the psychiatrists in charge. There are six grades, from *A* to *F*, the former applying to patients who have been sent home and the latter to patients who are mute and resistive, disturbed, not cooperative, etc.

[16]The same letter ratings were plotted against each other except for *D* cooperation which was plotted against *D* and *E* adjustment ratings. An examination of the definitions in both rating scales indicates that, on the whole, the adjustment ratings are more liberal to the patient than the corresponding cooperation ratings.

TABLE 12

INTERCORRELATION OF VARIOUS SOURCES OF VARIATION

	Educ.	Barr	Adj.			Irrit.
C.A.	*N*1: .00 (59)	*N*1: .02 (59)				
Educ.		*N*1: .85 (59)				
Coop.			*R*3I .41±.14	*R*3II .50±.12	*R*3III .32±.15	
Emot. det.						**R3I:** —.43±.18

31 better, four equal, one worse; for Period II—31 better, five equal, one worse; and for Period III—30 better, four equal, two worse. Although this is not conclusive evidence, it points to the great probability of a better attitude being shown to the experimental situation than in the daily activity of the patient.

Capacity has not as yet been defined. By capacity is meant what the individual is potentially capable of achieving after the removal of the more hidden effects of the psychosis. It refers to the level of achievement he could reach under optimal conditions. (Ability on the other hand, as pointed out earlier, is what he is able to do when the *immediate* interfering factors are eliminated.) In psychological work there seems to be good reason for accepting the achievement measures whenever one feels that they are not affected by the experimental situation, and are, therefore, good samples of the subject's customary behavior. It is perhaps a little artificial to treat cooperation as an external factor when dealing with a psychotic patient. His method of reacting to the experiment insofar as it is not a peculiar kind of reaction is part of his psychotic personality, and this is reflected to some extent in his score. Thus, to the extent to which tapping performance is an index of speed of movement and to the extent to which it is an index of general motor functioning, our scores might be used as measures of the state of the patient at the time as far as these characteristics are concerned, and used as a basis from which to work to determine the effect of various changes in the situation of the patient.

However, despite the advantages obtained from the scores, one is not satisfied with these descriptions of the patient. One is really interested in finding out what the ability of the individual is and also what his capacity is. Unfortunately, the latter is a rather difficult problem to solve and requires long, drawn-out procedures to determine. The first step in this direction is, of course, the determination of ability. If one finds that the patient comes up to the normal in his ability, one has good grounds for assuming that he also does in capacity. If he does not, then it is necessary to answer the question of the latter experimentally.

Since a considerable part of the argument about the determination of ability hangs on the rating system of cooperation, it is appropriate to spend a little more time on it. The classifications have

already been given. Obviously the judgment is a difficult one and is likely to be influenced by a number of factors—the major of which is the score. It might be argued that since the experimenter also rated the subjects, high attitude ratings were given to those with high scores and low ones to those with low scores. The following may be said with regard to the validity of the ratings: (1) They were all made by the same individual; thus there is a greater likelihood of stability in standards. (2) They correlated rather well with adjustment ratings—which is to some extent a check. (3) The scores distributed themselves quite normally about each cooperation rating, which is a rather strong statistical argument. (4) The same score is given to different ratings, i.e., a rating is not attached to a score. (5) The rating is not attached to a person—the same patient in different periods is given different ratings. (6) The ratings were made by an individual who had had considerable experience with them. These considerations lead one to believe that considerable faith may be placed in these measures. In any case it is the best measure available at present and the only one with which to work.

The data of the effect of attitude (called "cooperation" interchangeably) on performance[17] may now be presented to see whether the differences found can be accounted for on this basis. The table giving the distribution of ratings of the merged group by diagnostic type (Table 13) is first presented.

TABLE 13

INCIDENCE OF COOPERATION BY DIAGNOSTIC TYPES IN PERCENTAGE

Diagnostic type	Cooperation rating A	B	C	D	Not given
Hebephrenic (30)	7	33	50	7	3
Catatonic (12)	0	25	58	17	0
Paranoid (24)	33	58	4	0	4
Indeterminate (25)	4	56	40	0	0
Unclassified (24)	21	21	46	4	8
Simple (4)	0	75	25	0	0
Mixed (4)	0	75	0	0	25
All types merged	13	42	38	4	4

[17]By the standards used with the patients, all the normal subjects fell into the *A* group.

It will be noticed that for the whole group the great majority of the ratings fall in the *B* and *C* classes. The paranoids have the best ratings and the catatonics the poorest. This table is rather interesting in itself and deserves consideration for various reasons, but we cannot examine it further here.

Table 10 gives the correlations of cooperation with measures of central tendency and dispersion. It will be seen that there is a very marked consistency in the coefficients of mean score and rating. For the merged group (118) it is .66, for the large research group (79) it is .66, and for the three periods (36), .65, .65, .71, respectively. These are the highest correlations we have obtained from the sources of variation we are considering, and indicate a comparatively high degree of relationship between mean score and cooperation.

The correlations of cooperation with range for the three periods show interesting shifts over the three periods, —.16, —.47, .00, for the three respectively. (It is interesting to note that for the large group in the first period the coefficient is practically identical with that for the small group.) This suggests, in connection with the data already presented, that in the first period a great number of subjects—both good and bad cooperators—had high ranges, in the second period the poor cooperators were predominant among those with high ranges, and in the third period most of the subjects had low ranges. An examination of the scatter diagrams corroborates this interpretation, and would indicate that range is of some value as an index of attitude only at a certain stage of the tapping procedure.

It is of interest to point out something else with regard to the variation in its relation to cooperation. The standard deviation of the distribution of mean scores for the merged group is 5.58. That of the normals is 4.24. When cooperation is partialled out, the standard deviation for the patients becomes 4.21, identical with that of the normal group. The implication would be that the excess variation in the patient group is due to the cooperation factor.

It would be natural to say on the basis of the size of the correlations obtained and on the basis of the table presenting the cooperation ratings of the diagnostic types, that the correlation is to a large extent due to the fact that the poor performers, the catatonics, also have the poorest cooperation ratings, and the best

performers, the paranoids, also have the best cooperation ratings. This argument would be in line with those used in regard to age and duration of psychosis as factors. However, for various reasons, one is reluctant to argue thus. The variables mentioned before seem extraneous and independent as much from a psychological as from a statistical standpoint, whereas attitude, though statistically independent, seems, from the psychological standpoint, to be very much a dependent variable. One has the feeling that something more is involved than a correlation due to differences in cooperation of the various types.

This is corroborated by the following table.

TABLE 14

MEANS OF SCORES FOR EACH TYPE FOR EACH COOPERATION RATING

(Merged group)

	Par.	Heb.	Cat.	Indet.	Uncl.
A	27.5	26.0	——	19.0	26.5
B	22.0	22.0	13.0	19.0	21.0
C	22.5	18.0	15.0	19.0	15.5
D	——	14.0	11.0	——	12.0

It will be seen that even with cooperation kept constant—that is, comparing within the same cooperation levels—the catatonics consistently get the poorest scores, whereas the paranoids get the highest. Comparing within the type, cooperation seems to have a definite effect on the hebephrenic, unclassified, and paranoid groups, but none on the indeterminate group. The scatter diagram corroborates the implications of the table.

But even though it does not seem reasonable that cooperation is the fundamental factor which accounts for the difference between the types — especially between the catatonics and paranoids — still, from the standpoint of our argument with regard to ability, it would be wise to take only those cases of patients who meet the standards of attitude found in normals. For this purpose the 16 patients (from the merged group) who had *A* ratings were taken. Their mean is 25.1, while that of the normal group is 28.9. Although the normals' score is still higher, the difference is probably no longer significant—the first time that this has happened in the comparison of the two groups

as to means. This may be partially due to the fact that there are no catatonics in the group of 16 patients (8 *P*, 5 *U*, 2 *H*, 1 *I*).

When an analysis is made of the matched group of patients (n=26) on the basis of cooperation, it is found that there is practically no difference between the patients and their paired normals at the *A* level of cooperation, but with decreasing cooperation ratings, increasing differences between the two groups appear. Of course, there are only a few cases at each cooperation level, and these data should be considered with this in mind, but the consistency of the tendency and its corroboration of the results mentioned above strengthen the probability of the validity of the findings.

On the basis of these results it seems legitimate to generalize that to a large extent differences in scores from those of normals is due to attitude and that when attitude is kept constant the differences tend to vanish. Within the schizophrenic group there seems to be some difference between the types—a distinct difference—which is not covered by the attitude rating. The more detailed discussion of the rôle of attitude is postponed for another occasion.

We proceed to discuss various factors connected with the psychosis which may play a part in determining the scores obtained by the patients. There are two groups of phenomena with which one has to deal. In the first place there is the psychiatrist's rating which is based on the psychotic condition, already discussed.

An examination of Table 10 shows fairly constant correlations with mean scores of .34, .48, and .38 respectively for the three periods. A rather interesting point comes out in the correlations of adjustment type and range. The correlations are —.25, —.52, and —.06, for the three periods, which would seem to indicate that the stage of development of the tapping procedure represented by the second-period range is relatively the most sensitive indicator of general adjustment level. One is struck by the parallelism of the correlations, for the same period, of range and cooperation, and it is on this parallelism that the argument is largely based rather than on the significance of the differences between the coefficients. The consistency cannot be accounted for on the basis of the correlations between adjustment type and cooperation for they are not so very different for the three periods: .41, .50, .32. The cooperation correlations have already been discussed.

TABLE 15

CORRELATION ON MEAN SCORE AND SYMPTOM RATINGS

Symptoms	Mean score *R*3I	Mean score *R*3II	Mean score *R*3III
Initial psychomotor retardation	—.45±.13 (38)	—.46±.13 (38)	—.44±.14 (37)
Executive psychomotor retardation	—.44±.13 (38)	—.38±.14 (38)	—.38±.14 (37)
Reduction in volitional control	—.46±.14 (33)	—.42±.15 (30)	—.49±.14 (29)
Emotional deterioration	—.26±.16 (34)	—.25±.16 (34)	—.50±.13 (35)
Fatigability	.17±.23 (19)	—.10±.19 (29)	—.24±.18 (29)
Irritability	.52±.12 (37)	.45±.14 (36)	.14±.16 (38)

Note: N is given in parentheses.

During the course of the study a mental status rating scale was developed for use with research patients. It was used in rating on a five-point scale various symptoms and traits shown by the patients. The ratings were made on the wards by the psychiatrists, and, of course, independent of the psychological laboratory.

The correlations of mean score with some of these traits is presented in Table 15. For the items: "initial psychomotor retardation," "executive psychomotor retardation," and "reduction in volitional control," high correlations for this type of material and striking consistency over the three periods are found. The correlations on "emotional deterioration" are especially interesting. They remain constant for the first two periods, fairly low, —.26 and —.25 respectively, then increase to —.50 in the last period, which would seem to indicate that only when the individual had reached a degree of stability in his score did the test result serve as some index of emotional deterioration. The correlations on fatigability are interesting too. The low correlations are gratifying, since the test was definitely intended to eliminate the fatigue factor as much as possible. If the judgments of the psychiatrists are correct, and, in addition, test fatigability is a good indication of general fatigability, then the test is avoiding one of the difficulties it was intended to avoid.

The correlations with "irritability" are worth attention in connection with what was said about emotional deterioration and are very consistent with the findings on the latter score. The correlations are fairly high in the first two periods (.52 and .45 respectively) and drop to .14 in the last period. Is that not what one might expect? In some respects irritability is the obverse of emotional deterioration. The dictionary which is used with the rating scale defines "emotional deterioration" as "impairment—usually progressive—of ability to show normal emotional responses." If one looked into the scoring of the psychiatrists, one would probably find that "emotional deterioration" correlated very highly with "reduction of affect—amount" defined as "lack of manifestation of interest or feeling." This is the opposite of "irritability" defined as "abnormal responsiveness to stimuli." (The correlation of emotional deterioration and irritability is —.43.) It would mean that, in the early stages of the tapping experiment before a level has been reached, the score is a relatively good indicator of irritability but a poor one of emotional

deterioration, but in the later stage—when a limit has been reached—it is a relatively good indicator of emotional deterioration, but a poor one of irritability.[18]

6. *Other Aspects.* Considerable data are at hand on the intercorrelations between the tapping results and results on numerous other studies with the same subjects. It is deemed more desirable, however, to reserve these data for separate publication. It is also hoped that it will be possible to present at the same time an analysis of individual cases in relation to their psychosis.

Discussion

It is not our intention to enter into a lengthy discussion of the implications of the results since such material more logically belongs in the final paper of the contemplated series. However, there are a number of points connected rather directly with the present experiment which would bear consideration.

The situation with regard to variation is of interest. The schizophrenics show a greater amount of group variation and individual variation than do normals. There is less consistency of performance from trial to trial than is found in normals. Even in the manic group, which is much closer to the normal group in mean, we find that the variation is at approximately the level of the schizophrenics. This would lead us to believe that at least in the earlier stages of psychosis, the mere fact of psychosis makes for greater variability. The group of patients with greater psychosis duration—the non-research group—shows considerably less individual variation and somewhat less group variation.

Within the limits of the experiment the patients show improvement from period to period whereas the normals do not, despite the fact that the function is one in which there is little practice effect. It may be deduced from this that we are dealing with a slow adjustment of the patient to the experimental situation, that the achievement is not a good measure of ability and, of course, an even worse one of capacity; and that, at least in schizophrenics, it is necessary to be very cautious in judging ability (and, even more so, capacity)

[18]Cf. Strong (9, p. 79) and the references he gives. Tapping rate is discussed in relation to excitement and irritability.

from a first examination. We shall reserve a more detailed discussion of this point for a later occasion.

A related datum is the fact that when absolute values are disregarded, i.e., in the consideration of the place of appearance of best trial, the normals get the greater percentage of best trials appearing among the first five trials. It might also be pointed out that in the case of the trials giving the best scores there is relatively less improvement between these than there is between means when compared by periods within the group or when compared with normals. This would fit in with the implication of the above, that at intervals the schizophrenic may reach a level much nearer ability or capacity, and that these less frequently come early in a procedure, for there the variability is greater.

The results of this experiment give some evidence for accepting type diagnosis as legitimate when this is done conservatively. An examination of the types reveals differences in the profiles. On the one hand we have the paranoids who are most like the normals[19]—they have the best score, remain consistent over three periods, and show the best cooperation ratings. The catatonics, on the other hand, have scores which place them in a different universe as far as the distribution of the means of the first period are concerned. They show little consistency in the three periods, they have the most marked improvement over this time, and they also show the poorest cooperation ratings. Whereas the hebephrenic and indeterminate groups seem to resemble each other in many respects, e.g., in mean scores, in their consistency over three periods, and in their cooperation levels, there are a few characteristics which distinguish them sharply. The variation of the indeterminate patients is the lowest of all the groups, that of the hebephrenics is the highest; the scores of the indeterminate subjects do not seem to be affected by the cooperation level, those of the hebephrenics are highly affected. Both the differences and the similarities seem reasonable. In the case of hebephrenics one is dealing with relatively deteriorated individuals who are quite varied and in whom the psychosis is still active. In the case of indeterminate

[19]This greatest similarity to the normals has been found in most of the psychological and many physiological studies made at our hospital. It has also been pointed out by others.

patients one is also dealing with deteriorated individuals but with ones in whom the active aspects of the psychosis have died out, persons who have adjusted at a rather fixed level, and who are not much concerned with, or affected by, environmental variations. The effort they expend and the interest they take in the environment is low. Although there may be signs of differences in attitude which may affect the judgment of an observer, still these do not signify real differences in the level of interest and effort, and consequently in score. By definition the unclassified group is heterogeneous and one should therefore expect a medium level of mean score, great variation, medium cooperation, and medium effect of cooperation on score. These one does find. The fact that in the third period there seems to be a tendency for the groups to come together would argue for the differences between the types as being one of ability rather than capacity. On the whole, from the differences brought out it seems reasonable to argue that this experiment would adduce evidence on the side of those who argue for distinguishing between diagnostic types.

Summary and Conclusions

1. As part of an investigation of schizophrenia, a study of speed of movement was made with a modified alternate key tapping apparatus.

2. The groups included 123 schizophrenic, 60 normal, and 13 manic-depressive subjects. Parts of these groups had one or more repeated sessions at intervals of approximately three months.

3. The schizophrenics gave significantly lower mean scores and showed higher variation than did the normals.

4. When the schizophrenics are compared over three sessions they are found to have shown improvement from session to session whereas the normals, on whom data are available for two sessions, did not show any change.

5. The schizophrenics became less variable in succeeding sessions.

6. The data on the separate trials, of which there were ten in each session, corroborate the findings on the means.

7. Intra-session correlations indicate very high reliability and inter-session correlations fairly high reliability.

8. An analysis was made of selective factors, factors related to

the immediate experimental situation, and factors related to the psychotic condition to determine the part these may have played in determining the differences.

9. Age and duration of psychosis were found to have little effect on score. Whatever effect was present is probably accounted for by unavoidable selective factors due to diagnostic type differences.

10. Educational and occupational (Barr) levels seemed also to play little part in determining the differences.

11. The time of day of the experiment did not appear to affect the score.

12. The question of the effect of the attitude or cooperation is discussed and it is concluded that the major part of the differences between normal and schizophrenic subjects were due to this factor. The differences in mean score and in variation tended to disappear when attitude was taken into account.

13. The scores were related to various aspects of the psychiatric condition such as retardation, emotional deterioration, irritability, etc.

14. The diagnostic types of schizophrenia distinguished themselves in various ways. It is felt that the present experiment presents some evidence in the direction of indicating the validity of type differentiation.[20]

References

1. Benedict, F. G., Miles, W. R., Roth, P., & Smith, H. M. Human vitality and efficiency under prolonged restrictive diet. Washington: Carnegie Instit., 1919. Pp. 701.

1*a*. Boring, E. G. Learning in dementia praecox. *Psychol. Rev., Monog. Suppl.*, 1913, **15**, No. 63. Pp. iv+101.

2. Dunlap, K. Improved forms of steadiness tester and tapping plate. *J. Exper. Psychol*, 1921, **4**, 430-433.

3. Erickson, M. H., & Hoskins, R. G. Grading of patients in mental hospitals as a therapeutic measure. *Amer. J. Psychiat.*, 1931, **11**, 103-109.

4. Fisher, R. A. Statistical methods for research workers. (4th ed.) London: Oliver & Boyd. 1932. Pp. 307.

5. Gatewood, L. C. An experimental study of dementia praecox. *Psychol. Monog.*, 1909, **11**. Pp. 71.

---

[21]The authors are exceedingly grateful to Mrs. O. S. Baum and Miss C. Hall of the Biometrics Department for the enormous labor expended on the statistical analysis of the data.

6. HOSKINS, R. G., *et al.* A cooperative research in schizophrenia. *Arch. Neur. & Psychiat.*, 1933, **30**, 388-401.
7. JOHNSON, B. Changes in muscular tension in coordinated hand movements. *J. Exper. Psychol.*, 1928, **11**, 329-341.
8. REAM, M. J. The tapping test: A measure of motility. *Psychol. Monog.*, 1922, **31**, 292-319.
9. STRONG, E. K., JR. A comparison between experimental data and clinical results in manic-depressive insanity. *Amer. J. Psychol.*, 1913, **24**, 66-98.
10. "STUDENT." The probable error of a mean. *Biometrika*, 1908, **6**, 1-25.
11. TERMAN, L. M. Genetic studies of genius, Vol. I. Mental and physical traits of a thousand gifted children. Stanford Univ., Calif.: Stanford Univ. Press. 1925. Pp. xv+648.
12. WELLS, F. L. Motor retardation as a manic-depressive symptom. *Amer. J. Insan.*, 1909, **6**, 1-52.
13. WELLS, F. L. Studies in retardation as given in the fatigue phenomena of the tapping test. *Amer. J. Psychol.*, 1909, **20**, 38-59.

Worcester State Hospital
Worcester, Massachusetts

ÉTUDE DES FONCTIONS MOTRICES DANS LA SCHIZOPHRÉNIE: I. LA VITESSE DE TAPER

(Résumé)

Comme partie d'une investigation de la schizophrénie on a fait une étude de la vitesse du mouvement avec un appareil modifié de taper à clefs alternées. Les groupes ont compris 123 sujets schizophrènes, 60 normaux, et 13 maniaco-dépressifs, sur quelques-uns desquels on a répété l'expérience une fois ou plus à des intervalles de trois mois. Les schizophrènes ont donné des résultats moyens moins élevés d'une façon signifiante et ont montré une plus grande variation que les normaux. Le premier groupe a montré des progrès de séance tandis que le dernier n'a montré aucun changement. Les schizophrènes sont devenus aussi moins variables dans les séances suivantes. Les corrélations dans les séances et entre les séances indiquent une cohérence élevée. Une analyse faite des facteurs de sélection, des facteurs ayant relation à la situation expérimentale, et des facteurs ayant relation à la condition psychotique ont indiqué que l'âge, la durée de la psychose, et les niveaux d'éducation et d'occupation ont joué un très petit rôle. On a considéré que la plupart des différences ont été dues à l'attitude des schizophrènes. Les résultats ont montré une relation avec les divers aspects de la condition psychotique tels que la retardation, la détéreoration émotive, et l'irritabilité. Les types diagnostiques de la schizophrénie se sont distingués de diverses manières. On pense que cette expérience présente des évidences dans la direction d'indiquer la validité de la différentiation des types.

SHAKOW ET HUSTON

STUDIEN DER MOTORIK IN SCHIZOPHRENIE: I. SCHNELLIGKEIT DES KLOPFENS

(Referat)

Als Teil einer Untersuchung der Schizophrenie wurde eine Untersuchung der Schnelligkeit der Bewegung mit einem veränderten, abwechselnden Tastklopfenapparat ausgeführt. Die Gruppen bestanden aus 123 Schizophrenen, 60 Normalen, und 13 manisch-depressiven Patienten, bei denen einige das Experiment einmal oder mehrmals in Abständen von drei Monaten weiderholten. Die Schizophrenen ergaben bedeutsam niedrigere Durchschnittswerte und zeigten eine höhere Veränderlichkeit als die Normalen. Die erstere Gruppe zeigte eine Verbesserung von Zeit zu Zeit, während die letzteren keine Änderung aufwiesen. Die Schizophrenen wurden auch weniger veränderlich bei nachfolgendem Zusammentreffen. Intrasitzungs- und Intersitzungskorrelationen deuten auf eine hohe Zuverlässigkeit. Eine Analyse der selektiven Faktoren, der den experimentellen Situationen zugehörenden Faktoren, und der dem psychotischen Zustand zugehörenden Faktoren weist darauf hin, dass das Alter, die Dauer der Psychose, die Bildungs- und Berufsschichten keine grosse Rolle spielten. Der Hauptteil der Unterschiede waren die Folge der Einstellung der Schizophrenen. Die Werte bezogen sich auf verschiedene Seiten des psychiatrischen Zustands, wie Verzögerung, emotionelle Entartung, und Reizbarkeit. Diagnostische Typen der Schizophrenie unterscheiden sich verschiedenerweise. Es wird gemeint, dass dieses Experiment einige Evidenz in die Hand gibt, dass die Typunterscheidung zuverlässig sei.

SHAKOW UND HUSTON

A TECHNIQUE FOR INVESTIGATING RETROACTIVE AND OTHER INHIBITORY EFECTS IN IMMEDIATE MEMORY*

From the Psychological Laboratory of Smith College

JAMES J. GIBSON AND GERTRUDE RAFFEL

If a relatively meaningless geometrical form is shown briefly to a subject and, after an interval, he is asked to reproduce it, the reproduction is usually fairly adequate. But if the presented form is followed by several others in a series its reproduction, together with the succeeding ones, is typically very much less accurate, or may even be impossible. The subjects frequently report that succeeding forms "blot out" or "obliterate" the preceding forms (3). Occasionally a particularly distinctive form is reported to inhibit the clear perception of the one following it, especially if the latter is similar to it. Such facts, when taken in connection with evidence from memory-span studies (2), would indicate that various possible types of inhibitory effects can profitably be studied in the field of immediate memory and by the method of serial presentation of visual forms.

THE EXPERIMENT

In the following experiment we set out to measure the degree of retention of a visually perceived form as a function of its position in the series, or, in other words, the effect on its retention of being preceded or followed by other forms. A series of five figures, drawn on cards, was presented to each subject, at the rate of one every two seconds. The subject was instructed to look carefully at each one and to be prepared at the end of the series to draw *one* of the forms which would be designated by the experimenter at that time. These instructions were intended to produce an equal degree of attention for each position in the series. At the end, the form to be reproduced was designated, e.g. "the second," at a time just eight seconds after the instant when the subject last saw the form in question. In other

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, May 27, 1935.

words, the experimenter would say "the first" immediately after the last exposure but would wait four seconds before saying "the third." Consequently the time intervening between perception and reproduction was the same for each serial position.

In order to control all other factors affecting retention except position in the series five "crucial" forms were selected for reproduction and five series of forms were employed, one for each form to be reproduced. Twenty-five forms, in all, were shown to a subject, every subject going through the experiment described five times, once for each series, reproducing one of the crucial forms in each position. The subject was unaware of this plan, however, and had no way of guessing what serial position would be requested next. The position of any crucial form in its series was shifted regularly from one subject to another in such a way that any one crucial form appeared an equal number of times in each of the five positions. In the end, after 35 subjects had made 175 reproductions, each position in the series was represented by a group of drawings which was equated to the other groups with respect to differences among subjects and differences among the five forms used.

The forms used in this experiment were drawn on 2¼ by 3½ inch cards, within a central area of about one square inch on each card. Each form consisted of a contour of from four to six line-segments. No form was clearly suggestive of any familiar object or symbol or common geometrical shape; an attempt was made to select five crucial forms which were about equally meaningless. In addition to this set of 25 forms, another of 50 was made up for use in subsequent experiments. The latter forms were composed of six straight line-segments each.

This experimental technique was adopted on the assumption that the adequacy of the reproductions could be scored—that by inspection of the reproductions and the application of certain criteria, a quantity could be assigned to each which would serve as a measure of relative excellence of retention. If this were possible, we believed that it would provide an improved measure of retention over that obtained merely by calculating percentages of the number of forms reproduced (or omitted) at each serial position. A method of scoring reproductions would take into account not only number of forms recalled, but also the adequacy with which they had been retained.

Distinctions between varying degrees of adequacy of reproductions are not so difficult to make as might be supposed; we were able to establish sets of criteria by which the drawings could be scored with consistency and with considerable subjective certainty. It should be kept in mind that there were only five forms used but a large number of reproductions of each.

The scoring system was developed in the following way. The two experimenters set out independently to study the reproductions and draw up a list of criteria for scoring. The two methods were, by agreement, very different; the first (Method I) was relatively mechanical in its criteria and emphasized correctness of parts or details, the second (Method II) utilized somewhat more "subjective" criteria, and emphasized correctness of the form as a whole. The first gave a score of one point for each correct line in the reproduction, plus a point for correct orientation, one for correct relative proportion of parts, and one for correct size. The second gave a score of one if the general shape or impression of the reproduction was only vaguely similar to the original, two if the shape was somewhat similar, three if part was correct but part incorrect, four if all parts were correct but out of proportion, and five if the shape was approximately correct, plus a point for proper orientation and another for correct size. Both methods assigned a score of zero when the reproduction was wholly dissimilar to the original or when no reproduction could be given.

The reproductions were scored independently by these two methods. The correlation between the two sets was .64, which was only to be expected if the methods measured different aspects of retention, but both gave results which showed the same trend. A third intermediate method of scoring (Method III) was then developed on the basis of the other two. Scores were assigned on a basis of the correct component lines, as in I, but with additional criteria which gave more credit for general similarity of shape. The reproductions were rescored by this final method and it was found to give a correlation of .84 with the averages of scores by Methods I and II. The results will be given in terms of all three methods, but Method III is believed to be the most valid. This general technique for quantifying the adequacy of reproductions is probably not wholly satisfactory, but it shows the possibility of using nonsense forms as memory material for quantitative studies.

Results

The average scores for the five positions in the series, from first to last, were as follows. For Method I, 1.9, 2.7, 3.3, 3.9, 4.5; for Method II, 1.7, 2.1, 2.0, 2.7, 3.1; for Method III, 1.9, 2.5, 2.7, 3.3, 3.6. There is a progressive increase in retention-values as one passes from the first to the fifth serial position; only positions 2 and 3 by Method II show a reversal of this uniform progress. With large individual differences among subjects and among the forms, and with relatively few subjects, one cannot expect high indices of reliability. Nevertheless the difference between the averages of the first and fifth serial positions is statistically reliable for all three methods of scoring.

Any effect of the supposed factor of "primacy" in serial memorization is notably lacking in these results. In a series of impressed items where *S* does not know which item he will be called upon to reproduce and where the series is gone over only once, the first is evidently not a "favored" position. Our results are wholly in accord with those of Welch and Burnett (6) in indicating that the effect of primacy, as such, is an artifact.

Not only is this true but the supposed effect of "recency" in favoring later serial positions is present here, although the temporal factor in recency has been ruled out by the experimental conditions. It will be remembered that a reproduction was always called for eight seconds after the form in question had been exposed, whatever its position. Evidently the fact of better retention for later positions may not depend upon the *recency* of these positions at all, but rather upon the fact that they are *subsequent* positions. In our procedure, the first form was followed by four other perceptions, the last form by an empty interval, and others by proportional combinations of these interpolated conditions. The decrease in retention-values for the preceding items in the series must rest on some inhibitory effect of the succeeding impressions. This effect is obviously classifiable as retroactive inhibition, although we do not mean to imply by the use of this term that the effect is necessarily identical in all respects with that obtained when isolated tasks, themselves involving lists of items and repetitive memorization, are used.

Introspective reports on what happens to the memory-image during a serial presentation of forms bear out the above conclusion.

Two statements, by trained *S*'s, follow. "The first one seems to slip away as the next is perceived," especially "when a similar one follows." "First clearly seen. Certain I could have drawn accurately at this point," but when the next form appears, "first immediately gone." At the end of a series, the last form is often the only one represented by a clear image, but, either with effort or spontaneously, images of earlier forms may begin to reappear.

This retroactive inhibitory effect is, we may be sure, not the only detrimental influence at work in the course of the subject's task. If it were, the average retention-score of the last form in the series would be the same as that obtained when one form alone is exposed and reproduction is required after the eight-second interval. Such is obviously not the case. When 28 subjects were shown a single form the reproductions showed a much higher average score (Method III) of 6.7, and the great majority of reproductions were essentially "correct." This average is to be compared with 3.6 for the last in a series of 5.

It would be easy to conclude that only a proactive inhibition (1) need be postulated in order to explain this decrease in score for the last of five forms as compared to one alone. Memory images of preceding forms may increasingly inhibit the clear perception of each succeeding one and with such a proactive effect superposed upon the retroactive effect, the facts can be harmonized. Presumably the total effect of the retroactive would outweigh the proactive inhibition. The additional postulate of a proactive inhibition, however, does not exhaust the possibilities of all the detrimental influences which may be at work. A priori, a form apprehended singly, and a form apprehended in a series, differ in at least the following respects. In the first place, the form in a series bears the character of group-membership. Besides being a *form,* it is also the fourth, the first, or the middle form of a temporal series. Under some conditions, this position-character may be vague or even wholly absent, but in our experiment it was important since the forms had to be designated for reproduction by their serial position. What effect this fact may have upon the retention of a serial form as compared with a single form, we do not know, but it may well be important. In the second place, the form in a series (or rather, the corresponding process of perception, retention, and reproduction) may be modified

by various types of interaction within the series. In so far as such interaction is detrimental to "accurate" reproduction, we may distinguish the following possible types.

1. A detrimental effect of later perceptions on the retention-processes, or traces, of previous perceptions.
2. A detrimental effect of traces on previously established traces.[1]
3. A detrimental effect of existing traces on perceptions.
4. A detrimental effect of previously established traces on subsequently established ones.
5. A detrimental effect of traces earlier and later than the "selected" trace on the reproduction-process of that trace.

These theoretically possible types of effect are probably not completely isolable from one another, except logically. Neither, for that matter, are the effects of retroactive and proactive inhibition, in preference to which the above list is given. Nevertheless, it should be noted that effects 1 and 2 would produce a progressive increase in scores as one passes from the first to the last serial position while effects 3 and 4 would produce a progressive decrease. The fifth, as stated, is equivocal in this respect. Since the actual scores increase progressively, it may be concluded at least that the influence of effects 1 and 2 overweighs the influence of 3 and 4.

The foregoing results raise a number of problems and suggest several corresponding lines of experimentation. There is first the question of why the "primacy" effect usually appears in serial memorizations; i.e., to what factors it is due if not to primacy as such.[2] Then there is the question of whether recency as such favors retention in immediate memory apart from what we might call "subsequency." Since we have seen that subsequent items in a series are better retained quite independently of their temporal recency, it may be that this factor explains *all* of the so-called recency effect. This question is obviously of interest in that it raises the issue in primary memory which has already been brought up in

[1]D. M. Purdy (5) has recently quoted evidence to suggest that new impressions as such may not cause forgetting of old ones unless they set up traces.

[2]One of the writers (Raffel) has since performed an experiment on this problem to be published in the *Amer. J. Psychol.*, entitled "Two Determinants of the Effect of Primacy."

secondary memory (4): whether forgetting is a function of mere lapse of time, or of the activity filling the time. Finally, there is a third question suggested by our results: what is their relation to the fact of the memory span and to the retention of temporal patterns of impressions generally?

Some additional experimental evidence has been obtained which bears upon the second of these questions and has some implications for the third. We wished to know whether the length of the period between perception and reproduction had any significant effect on degree of retention, and also whether retention is still a function of subsequency in a shorter series of impressions. Series of three forms, instead of five, were presented to subjects using the same method as before. There were three crucial forms, and each subject was given three series, making three reproductions, one for each serial position. All influences except that of serial position were equated. Three groups of 24 subjects each performed the experiment, with three different constant intervals between perception and reproduction of four, eight, and twelve seconds. Each group yielded 72 reproductions.

The results demonstrated that, within the range of these time-intervals, recency has little significant effect on degree of immediate retention. The average scores by Method III for all reproductions at each of the three time-intervals were 3.86, 3.75, and 3.41 for the four-, eight-, and 12-second intervals respectively. Other comparisons gave approximately the same results. The differences are in the expected direction, but are slight as compared with the differences produced by subsequency in the last experiment.

TABLE 1

SUMMARY OF RESULTS IN TERMS OF AVERAGE SCORES FOR EACH SERIAL POSITION

No. subjects	No. forms	Time-interval	Scores for positions, first to last					Av. for all positions
35	5	8 sec.	1.9	2.5	2.7	3.3	3.6	2.80
24	3	4 sec.			4.0	3.4	4.1	3.86
24	3	8 sec.			3.8	3.6	3.9	3.75
24	3	12 sec.			3.4	4.0	2.8	3.41
		All 3 groups			3.72	3.68	3.62	3.67
28	1	8 sec.					6.7	6.7

The averages for the three serial positions give a less expected outcome. No "subsequency-effect" whatever appears when the series is limited to three forms. The average scores from first to third positions are 3.72, 3.68, and 3.62. Table 1 shows these results, together with a summary of the preceding ones. The reason why a shorter series shows no differences between serial positions while a longer series does is not assignable with confidence. But it is probable that the complex of inhibitory influences which we may suppose to be at work among the items of a series of impressions is different in its total effect when the series is more easily grouped or unified by reason of being short than when not so easily unified. The situation is probably similar to a series of items below the memory span as contrasted with one at or above the span.

Summary and Conclusions

1. The retention of geometrical nonsense forms, presented in a series, is an inverse function of their serial position when five forms are given, but does not vary with position when three forms are given.

2. No effect of "primacy" obtains in these results. One therefore must go behind the fact of primacy as such in explaining the usual effect obtained in serial memorization.

3. The effect of "recency" has been shown to depend principally on subsequency of serial position rather than on recency as such.

4. Instead of the laws of primacy and recency, future investigations should bear on the complex of possible inhibitory influences which may be at work among the items of a series of impressions and on their "position character" and membership in a temporal group.

References

1. Foucault, M. Les inhibitions internes de fixation. *Année psychol.*, 1928, **29**, 92-112.
2. Gates, A. I. The mnemonic span for visual and auditory digits. *J. Exper. Psychol.*, 1916, **1**, 393-403.
3. Gibson, J. J. The reproductions of visually perceived forms. *J. Exper. Psychol.*, 1929, **12**, 26.
4. McGeoch, J. A. Forgetting and the law of disuse. *Psychol. Rev.*, 1932, **39**, 352-370.

5. PURDY, D. M. The theory of forgetting. *Amer. J. Psychol.*, 1934, **46**, 339-340.
6. WELCH, G. B., & BURNETT, C. T. Is primacy a factor in association-formation? *Amer. J. Psychol.*, 1924, **35**, 396-401.

Smith College
Northampton, Massachusetts

UNE TECHNIQUE POUR L'INVESTIGATION LES EFFETS RÉTROACTIFS ET D'AUTRES EFFETS INHIBITEURS DANS LA MÉMOIRE IMMÉDIATE

(Résumé)

Une forme non-sens quand vue seule est beaucoup mieux rappelée et reproduite que lorsqu'elle est précédée ou suivie d'autres formes dans une série. On décrit une méthode pour mesurer la rétention d'une forme visuellement perçue comme fonction de sa position en série, indépendamment du facteur récence. Dans une série de cinq formes il y a une augmentation progressive des résultats de la rétention depuis la première position jusqu'à la cinquième. Il n'y a aucun effet de primauté. D'ailleurs, la valeur supérieure pour la rétention des positions subséquentes doit se baser sur cet "état subséquent," non pas sur la récence temporelle.

Une autre expérience, employant des séries de 3 formes au lieu de 5 et 3 divers intervalles de temps entre la perception et le rappel, montre en premier lieu que la récence telle quelle a un effet négligeable sur le degré de rétention, et en second lieu que la rétention n'est plus une fonction inverse de la position en série quand seulement 3 formes sont données dans la série.

On croit qu'à l'avenir on doit faire attention non pas aux "lois" de primauté et de récence dans la mémorisation en série, mais aux divers types d'effets inhibiteurs qui peuvent être engagés, et à la qualité de membre des parties individuelles dans un groupe temporel.

GIBSON ET RAFFEL

EINE METHODE ZUR UNTERSUCHUNG DER RÜCKWIRKENDEN UND ANDEREN HEMMENDEN WIRKUNGEN AUF DAS UNMITTELBARE GEDÄCHTNIS

(Referat)

Wenn eine sinnlose Form allein wahrgenommen wird, wird sie viel besser behalten und wiedergegeben, wenn andere Formen in einer Reihe ihr vorangehen oder folgen. Eine Methode zur Messung der Behaltung einer visuell wahrgenommenen Form wird als eine Funktion ihrer Reihenstellung beschrieben, ganz unabhängig vom Neuheitsfaktor. In einer Reihe von 5 Formen gibt es eine fortschreitende Vermehrung der Behaltungswerte von der ersten bis zur fünften Stellung. Es gibt keine Wirkung des Vorranges. Ferner muss der überlegende Behaltungswert der nachfolgenden

Stellungen auf diesem "Nachfolgen" beruhen, nicht auf der zeitlichen Neuheit.

Ein weiteres Experiment mit einer Reihe von 3 Formen anstatt 5 und mit 3 verschiedenen Zeitabständen zwischen der Wahrnehmung und Wiedererrinnerung beweist, erstens dass die Neuheit als solche eine nicht zu beachtende Wirkung auf den Grad der Behaltung besitzt, und zweitens dass die Behaltung nicht mehr eine umgekehrte Funktion der Reihenstellung ist, wenn nur drei Formen in der Reihe gegeben werden.

Es wird geglaubt, dass die Aufmerksamkeit in der Zukunft nicht auf die "Gesetze" des Vorranges und der Neuheit in dem Reihenauswendiglernen, sondern auf die verschiedenen Arten der Hemmungswirkungen, die wirksam sein könnten, gelenkt werden sollte, und auf die Mitgliedschaft der Einzelheiten in einer zeitlichen Gruppe.

GIBSON UND RAFFEL

ON THE FUNCTIONAL RELATIONS OF THE SENSE ORGANS TO ONE ANOTHER AND TO THE ORGANISM AS A WHOLE*[1]

WALTER BÖRNSTEIN

I. PSYCHOLOGICAL PART

One of the most essential achievements of sense-physiology during the past century was the demarcation of the domains of the single senses from one another by the "law of specific energies of the sense-nerves" formulated by Johannes Müller (19). On that law as a foundation Helmholtz (7) built up the doctrine: that between the domains of the single senses, the "modalities," there is no communication (*Übergang*). Nevertheless, the question as to whether there do exist intermodal relations—that is to say, functional connections—between the different sense organs has since then again and again been discussed.

The proof that sight and hearing have one faculty in common, namely that of perceiving *"brightness* and *darkness,"* has been furnished by Mach (14), Brentano (3), and Wolfgang Köhler (11); and recently it has been shown by von Hornbostel (8) that this faculty is the basic function of all the sense organs.

I refer to the following passage from von Hornbostel's "Dreiecks-Versuch" (triangle-experiment):

> To a particular smell, say benzol, the correspondingly bright grey is chosen on the color-disc, and to the same smell, from a series of tuning-forks, the correspondingly bright tone. In a third equation, between this tone and the greys of the color-disc, the subject unconsciously chooses the same grey as in the first equation.

*Communicated by Karl S. Lashley, accepted by Carl Murchison of the Editorial Board, and received in the Editorial Office, May 28, 1935.

[1]This thesis is offered by the author as a survey of experiments so far made, and as a program of their further prosecution. It is intended, as soon as circumstances allow, to furnish experimental proof of the theories put forward, and to give an account of them in monograph form. Cf. Börnstein (1, 2).

This proves that the degrees of brightness in the optical, acoustic, and olfactory fields can be directly compared with one another. The same applies also to the other fields of sense.[2]

From this it follows: *that the faculty of perceiving bright and dark is a function common to the fields of sense.* To this there can correspond in the bodily field only an *unspecified* process.

The degree of brightness in the optic and acoustic fields depends upon the amplitude and frequency of the light- or sound-vibrations: yellow (amplitude being equal) is brighter than red; a high note is brighter than a deep one. In the same way, as von Hornbostel (8) has shown, in other fields of sense the *brightness* of a perception-content is plainly determined by the nature of the stimulus. Sharp skin stimuli, for instance, have a bright effect, blunt ones a dark; cold has a bright effect, warmth a dark, urea and cane-sugar taste bright (cold), magnesium sulphate dark (warm);[3] a sharp pain is bright, a dull pain dark.

For *smells,* von Hornbostel in 1931 drew up the following series of increasing brightness by way of example: xylenol 1, 3, 5; *a* cresol; benzol; *a* pine; citral; citronellal; citronellol d.; citronellol l.

According to von Hornbostel, the faculty of perceiving brightness and darkness is *an intermodal primitive function of all sense organs.* If this theory is correct, it is to be expected that a brightness-excitation taking place in *one* field of sense will be regularly modified—strengthened or weakened—by a supervening stimulus from another field of sense.

Whether and in what direction such an *intermodal reciprocal influence* exists, is now to be stated.

As test organ for my experiments I have utilized *the sense of sight,* as changes in the phenomena of brightness are most easily observed in that field.

From a larger complex of investigations, I have here chosen *three series of experiments.*

---

[2]This problem was recently summed up by L. Székély (26).
[3]Cf. the essay of von Hornbostel's pupil Michael (18).

First Series of Experiments.

The principle of the additive effect of brightness excitation.

Mode of conducting the experiment: The subject sits in darkness, about two meters distant from a dull-surfaced disc of 35:25 cm. lit evenly from above the threshold. As the disc is looked at, there occur, with differing strength in different individuals, spontaneous variations in the subjective brightness of the disc, which are first of all observed as to their degrees. In ten to fifteen minutes from the start of the experiment, while the subject continues to contemplate the disc, I bring into play non-optical bright or dark stimuli.

It is seen that when bright stimuli are brought to bear, for instance harmonium notes from the second to the fourth octave,[4] or a bright smell, taste, or tactile stimulus, the disc often lights up subjectively far above the spontaneous variations, that its contours become more sharply defined, that it appears more solid; and that when dark stimuli are brought to bear—deeper notes, a dark smell or taste, then the disc darkens, its contours become blurred, and the surface takes on a looser appearance.

In a few cases it was possible by means of dark stimuli to cause the spontaneously perceived disc *to vanish completely*: "the disc darkened until it was no longer there," said one subject when four deep notes were struck; and "it appears uncommonly bright," when four high notes were sounded. In another case, an objective diffused spot of light of about one-and-a-half inches in diameter, that happened to be visible close to the disc, was observed by the subject to disappear at the deep notes; while at the high notes it was remarked that "it shines with extraordinary brightness, oscillates, dances up and down like a roundabout." For a third subject, to whom were presented on a disc a number of concentric luminous rings of different grades of brightness, the sounding of high notes had the effect of bringing within the threshold of notice one of the rings which had not previously been perceived. At the same time it was remarked: "my eyes are moving like mad."

[4]In the case of harmonium notes this effect usually occurs only when three or four notes have been struck simultaneously (whether a discord or a harmony is immaterial).

The same result was obtained on presenting bright and dark smell and taste stimuli.

From this first series of experiments it can be seen that: *bright non-optical stimuli have upon an optical brightness-perception a brightening effect; and dark stimuli a darkening effect.* Heteromodal supervening sense-stimulation affects the already active brightness-excitation of the optical sense; and that not by contrast—i.e., in a weakening manner—but in the direction of its own "brightness," thus giving the *principle of the additive effect of brightness-excitations.*[5]

This result I have obtained in 35 out of 50 cases upon which I have experimented. Fifteen subjects, that is to say *30 per cent of the cases, showed no influencing* of the optical brightness-excitation by non-optical sense stimuli. I shall revert to this fact later (see below).

Second Series of Experiments.

The heightening of bodily tone ("Körpertonus") by brightness-excitations; "phosphenes" as an expression of heightened tone. From the reports of the experiments adduced above we have seen that some of the subjects, in addition to strengthening of the optical *brightness-excitation,* were conscious also of an *increase of movement* owing to the brightness-stimuli—the movements being either ascribed to the object under observation or else experienced in the subject's own body. From the investigations which have been conducted by Ladislaus von Stein (25) for 25 years past, and from more recent investigations [E. Metzger (17) among others], we know that *optical* brightness-excitations always *heighten bodily tone.* An equal conformity to law, as has been shown by earlier experiments of my own,[6] is characteristic also of *non-optical* brightness stimuli. That is to say: *brightness-exciting stimuli of every kind heighten bodily tone.*

From this it follows that: there exists an intimate functional connection between the brightness-excitations taking place in the sense

[5]Here I would draw attention to researches made by K. Zietz (27) and P. v. Schiller (22).

[6]Concerning my observations made in the "tone-apparatus" upon a patient with disinhibitions (Enthemmungen), Goldstein has reported in the *"Schweiz. Arch. f. Psychol. u. Neur."*

organs and the functional state of the *muscles* of the body. The nature of this connection will become clearer if we glance for a moment at the *functional mechanism of the retina.*

The retina reacts to brightness- and darkness-exciting stimuli in a twofold manner: by the adjustment of the flow of pigment, and by changes in the state of contraction of its cones. On exposure to light the pigment of the retina flows down between the rods and cones, and the cones contract (*"light-adjustment"* of the retina). On exposure to darkness these processes are reversed (*"dark-adjustment"*).

Now it is well known that heightenings of muscular tone at any one point in the body heighten tone in the rest of the body. It is these facts which furnish the starting-point for my working hypothesis upon which my *second series of experiments* is based.

Working Hypothesis: That the contractile substance of the retinal cones is functionally equivalent to the muscular system of the body, and forms a part of the tone-apparatus.

If this hypothesis be correct, then *heightenings of the bodily tone* —whatever may be the causes producing them—must always effect an increase in the contraction of the retinal cones, that is, *"light-adjustment."* In the expectation that light-adjustment would make itself subjectively noticeable by the occurrence of "phosphenes," that is to say, entoptical (colored or colorless) light-phenomena, I carried out the following experiments:

1. *Heightening of bodily tone by active contractions of muscles of the body:*

> The subject lies relaxed in darkness, until the eyes have completely adapted themselves to the darkness: that is to say, until it has become quite dark in front of the eyes, or until a motionless diffused field of grey presents itself.[7]

If now a powerful contraction is effected at any part of the body, say for instance an arm is strongly flexed at the elbow or extended against passive resistance, then there become visible to the eyes—caused by the movement in question—luminous phenomena, which,

[7]The interval of time elapsing before the "dark-adaptation" took place varied considerably; from a few minutes to three quarters of an hour. In the case of some subjects it had not taken place even then.

moreover, are stronger on that side of the body upon which the contraction takes place.

2. *Heightening the bodily tone by non-optical brightness-exciting stimuli:*

Again the subject lies completely at rest in darkness, until adaptation has taken place. Then bright and dark non-optical stimuli (sounds, smells, tastes, or tactile stimuli) are brought to bear upon the subject.

These experiments I carried out upon 70 persons. The result was clear. Two instances may suffice:

> *Instance No. 1.* (The subject experiences first of all progressive adaptation to the darkness): At 15.30 at the beginning of the test, the subject says: "I see streaks in movement . . . a streak running downward . . . waves like when a stone falls into water . . . a dark background, with lighter waves and lines upon it. Somewhat on the move, rather like an ant-heap, a swarming without any special order."—15.45: "Now it is all a grey crepe, with very close meshes, very *homogeneous.* When I am speaking, whirls appear in this mist. . ."[8]—15.55: *A bright smell* (citronellol l) is presented: "Oh! How beautiful! *Very bright! Very bright!* It shines like the Milky Way, little dots like needle-points! . . . Now they are disappearing." —At 15.58 a *very dark smell* (xylenol) is presented: "A *slight darkening* and breaking of the homogeneity. . . . Now it is the same as before."—A smell less bright than citronellol l—citral —makes fewer brightness phosphenes. A smell less dark than xylenol—cresol—darkens still less, and for a shorter time, than xylenol.
>
> *Instance No. 2:* Another subject, just after being placed in darkness, saw colored rings which, rhythmically moving, gradually diminished in size. After a few minutes the rings disintegrated and the color gradually paled away to a diffused brightish grey; their movements, on the other hand, gradually grew slower, and at last ceased altogether. The subject (now adapted to the darkness) had only a "slight impression that there is still something there—no longer anything optical." Now a *bright smell* was presented: after a few seconds colored dots appeared, which grew bigger and bigger and executed movements. The subject said in astonishment: "Curious! Suddenly I have a field of vision again!" The bright-phosphenes

[8]Speech is muscular work for the subject.

> may generally be specially easily and practically induced by introducing into the mouth gustatory matter, such as urea or cane sugar, and also by brushing or scratching the skin.

Many subjects when presented with bright smells, tastes, etc., perceive bright-colored, usually greenish-yellow, or also bluish phosphenes; when subjection to dark stimuli takes place, the phosphenes are mostly colorless-dark or brown.[9]

Generally speaking, the following were the results obtained in the second series of experiments: *Heightening of bodily tone of every kind*—caused by active contractions of muscles as well as by non-optical brightness-exciting stimuli—after the subject has become adapted to the darkness "brightness-phosphenes" ("Hell-Phosphene"), i.e., entoptic phenomena (not produced by outward optical stimuli) which are either colorless-light or colored, as already described, and mostly in lively motion. Entoptical phenomena produced by *dark stimuli* in the case of persons whose sight has become adapted to the darkness are in general colorless-dark, sometimes brown, and either at rest or only feebly moving; they occur more rarely, and usually more feebly.[10] We thus get: *induction of lightness-phosphenes by non-optical brightness-exciting stimuli, and of darkness-phenomena by non-optical darkness-exciting stimuli.*

The intensity with which these phosphenes occur varies from individual to individual. It shows all stages, from a plastic solidity—so that the phenomena are seen sharply defined, and at a certain distance projected outward—down to such a vagueness of impression that the subject does not know whether anything is *visible,* or, as it was formulated in one case, a *movement* is taking place inside of him.[11]

Here again, in about *30 per cent* of the cases no influencing of the entoptical phenomena by heteromodal sense-stimuli was observed; and yet they were the same subjects who had shown themselves refractory to such induction in the first series of experiments.

Appendix to the second series of experiments. The objection, that

---

[9]Some of these phosphenes, as painted by subjects, will be shown printed with the psychological monograph.

[10]Lightness-phosphenes still persisting are caused to disappear more quickly by darkness-exciting stimuli.

[11]The observations of these indefinite states appear to me to be of special importance in relation to the question of the *common root of optics and motorics.*

the results observed in these experiments might be founded upon "associations," is automatically refuted by the fact that phosphenes are also induced by brightness-exciting stimuli which are *imperceptible to the senses* of the subject. If for instance I cause ultra-violet rays (artificial sunlight) to play *upon the skin,* say of the back, of a subject after adaptation to the darkness, and whose head has been completely insulated against rays of light, then lightness-phosphenes appear, after a time of latency of half a minute to one and one-half minutes, at a distance of say four feet from the ray-source. Thus, one subject, upon the skin of whose back I caused rays to be played unknown to the subject, said: "It is becoming lighter . . . more and more . . . it is quite light!" Other subjects become aware of more or less luminous circumscribed round formations, or diffused colored mist or similar appearances. These lightness-phosphenes generally persist for a few seconds after the radiation has ceased. If the radiation is repeated on the same part of the skin, the phosphenes do not at first reappear.[12] In these investigations I have, in the case of 20 subjects, *always obtained a positive result. Heat rays,* however, *induce no* phosphenic effect whatever! Under favorable circumstances, lightness-phosphenes were observed when the skin of the subject was subject to visible light in the same way—mostly in the case of blue light, and hardly at all in the case of red light.

This inquiry into the question of phosphenes proves, therefore, to be a very fine test for such changes in the state of the body as are brought about by brightness-exciting stimuli in a person after adaptation to darkness.

Project for a Third Series of Experiments.

The starting point is the phenomenon observed by Purkinje: namely that the subjective brightness[13] of individual colors does *not diminish in the same degree* when the intensity of the light is reduced. For instance, a blue, which by daylight appears to be equal in brightness to a given shade of red, is in twilight seen to appear much brighter than the red.

---

[12]This is of importance for the theory; see below.

[13]For these experiments I found it most appropriate to present *taste-stimuli*: urea as a brightness-exciting stimulus, and magnesium sulphate as a darkness-exciting stimulus.

The experiment is carried out as follows: Two revolving colored discs are placed side by side, one of which always bears the same shade of red, and the other a blue color that can be rendered darker or lighter by a measurable admixture of black. To begin with, in the twilight, both colored discs are adjusted to equal brightness. Then non-optical brightness- and darkness-exciting stimuli are brought to bear, and a new brightness-equation established, in order to ascertain whether, and if so, in what direction, a modification of the Purkinje phenomenon has taken place. This method of carrying out the experiment has the advantage that it can be executed without the subject's seeing through it, and also that the results obtained can be determined *quantitatively*.

These experiments have so far been carried out only upon nine persons. Moreover, those subjects who had been refractory in my first series of experiments proved refractory here also; that is to say, heteromodal sense-stimuli in their case did not influence the optical brightness-excitation. Those subjects who had proved positive in the first series of experiments, except one case, reacted positively here also. This experimental material, however, is not yet sufficient to allow of drawing binding conclusions.

Refractory cases. With the exception of induction of light-phosphenes by ultra-violet stimuli (which has always given positive results), I have in all my experiments, in about *30 per cent* of the cases, and independently of the age of the subjects, obtained a negative result. This points to the possibility that there may exist connection inherent in the *structure of the organism as a whole.*

On the psychological side, the question of the interrelations between the structure of the organism as a whole and the mode of perception has been worked out preeminently by E. R. Jaensch (9).

Jaensch distinguishes between two types of psychic constitution: the *"integrated"* and the *"disintegrated"* type. By the integrated type, Jaensch understands, in contradistinction to the disintegrated type, persons in whom there exists an especially intimate functional connection between all sensual performances on the one hand, and the motorium on the other.

Now, in the Ambulatorium for Constitutional Medicine (Director: Private-dozent Dr. Walter Jaensch) I have investigated with Mandowsky (10, 15) so far 19 cases in such a manner that each of

us has applied his own methods independently of the other (15). At the conclusion of our experiments we compared results, and found, in 15 out of 19 cases, complete agreement in regard to the designation of the subjects as "positive" (Börnstein) or "integrated" (Jaensch-Mandowsky) and "negative" or "disintegrated" cases. These figures are not, it is true, sufficient to base a theory upon. They do, however, indicate that further investigation in this direction will be fruitful of results.

Relations to the clinic. My clinical observations also prove that the study of the reactions of the organism to brightness- and darkness-exciting stimuli is calculated to clear up problems which have *a scope far beyond the psychology of the perceptions,* and that here we have to do with *fundamental biological processes* of manifold significance for clinical practice as well.

The clinical investigations which I have so far made have had to do with (1) phosphenes caused by migraine and trigeminal neuralgia; (2) the relation between lateral displacement of the phosphenes and the seat of the malady in unilateral cerebral lesions; (3) the question as to whether there exists any connection between "dermographism" and phosphenic susceptibility; (4) the influencing of disturbed bodily tone (torticollis)[14] by non-optical brightness- or darkness-inducing stimuli; (5) with the occurrence of phosphenes, in case of visual-ray injury, in the hemianopic field of vision; (6) with the question of the local-diagnostic significance of phosphenes in brain maladies (differential diagnosis between supra- and subthalamic foci).

It would lead too far here to enter into the significance of the faculty of perceiving "brightness" and "darkness" in the biology of the *child* and of *animals.*[15]

II. Physiological Part

On the Physiological Substratum of the Functional Interrelations between the Sense Organs and the Organism as a Whole. Olfactory, auditory, tactile, and gustatory stimuli affect the optical and the bodily-motorial system primarily, though not exclusively, directly

[14]Cf. Quadfasel-Krayenbühl (21).
[15]Cf. F. T. Perkins (20), v. Schiller (24), etc.

by *neural paths*. Anatomy indicates that this functional connection is established by way of the *thalamus opticus*, as well as partly via the midbrain.[16]

On the other hand, to report upon how far the optical and bodily-motorial systems influence the non-optical sense-systems is not my present task.

It is known that the regulation of bodily tone occurs because of the intimate connection between the animal and the vegetative nervous system on the one hand, and the inner-secretory metabolism on the other. This fact indicates in itself that not only is the nervous system influenced by brightness- or darkness-inducing stimuli, but that there occurs a *"modification" (Umstimmung) of the whole organism*. The inner-secretory (hormonal) modifications we were already able to determine experimentally in the second series of investigations: when the skin is radiated the phosphenes cannot be set up along neural ways, but only by lawful changes in *that substance* which flows through all the organs: *the blood*. There must, therefore, be set in activity in the skin by the ultra-violet rays some substance which is conveyed through the organism by way of the blood, and which, by its action upon the optical system of a person who has become adjusted to the dark, can be perceived as "lightness-phosphenes."

Fourth Series of Experiments.[17] In the case of the frog I have been able to prove experimentally the *hormonal modification* as a consequence of simple exposure to light: Blood from a "dark" frog (i.e., one that had been kept in darkness for 12 to 24 hours) injected into the enucleated eye of another dark frog induces no change. *The blood of a frog that has been exposed to light induces lightness-adjustment in the retina of the eye of a dark frog.*

---

[16]The view here expressed, it may be noted, is strengthened by the experiments on rats reported by Lashley (13).

[17]The experiments on animals carried out by me were conducted at the Physiological Institute of the University of Frankfort-on-the-Main (Director Professor Bethe). As soon as I am in a position to do so I shall report upon them in detail and carry them further. I wish here to thank Professor Bethe for having placed the resources of his Institute at my disposal, and for having given me the benefit of his advice in regard to the difficult technical execution of the experiments, and also for having interested the "Notgemeinschaft Deutscher Wissenschaft" in the work. To the latter body I am also gratefully indebted for assistance in connection with the experiments.

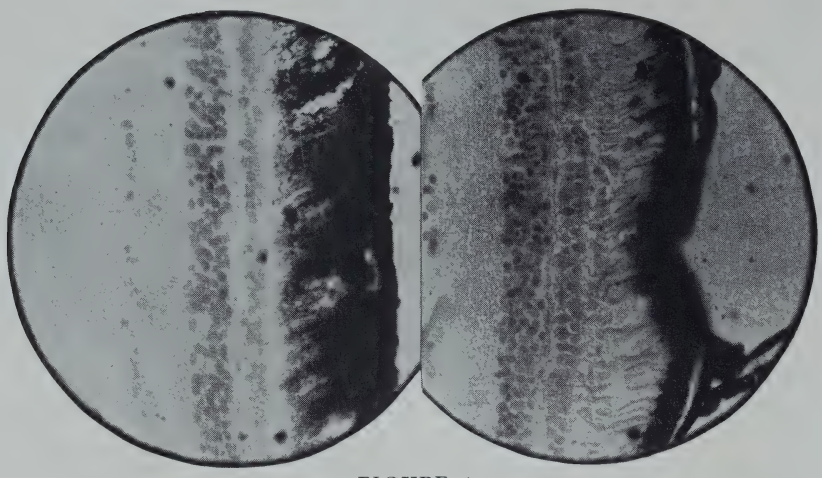


FIGURE 1

ENUCLEATED EYES OF THE DARK-ADAPTED FROG

1. Retina in position of adjustment to "dark."
2. Retina in position of adjustment to "light" caused by "light-blood."

This experiment proves that in the blood of the frog, in consequence of exposure to light, there is formed some substance which, transferred to the insulated "dark" eye of another frog, modifies its retina in the same manner as does expose to light of the eye itself. The eye of the frog thus under exposure to light acts as an *inner-secretory organ*[18] passing on into the blood a substance which develops a hormonal effect. The human skin, the "bodily-tone" apparatus, and all sense organs under the action of brightness-exciting stimuli evidently all behave in the same way.

Without wishing in any way to prejudge its nature, I have designated this substance, which I have ascertained to be produced in the case of the frog, *"bright-hormone" ("Hell-Hormon").*

Investigations conducted in other quarters with regard to the distribution, significance, and lability of acetylcholin strongly suggest that the *"bright-hormone"* in question has some relation to that substance, *being antagonistic to the acetylcholin.*

From the researches of Engelmann (5) and later investigators[19]

[18]Cf. Merkulow (16).
[19]Cf. Dittler (4).

we know that lightness-adjustment of the retina of an intact frog in darkness is caused not only by exposure of the eye to light, but also by other quite definite stimuli; such, for instance, as exposure of the skin to light, rubbing, and pinching of the skin, powerful muscular actions of the frog. The same stimuli induce in a human being adjusted to darkness, as reported above, the occurrence of lightness-phosphenes! It will be a task for further investigation to observe in general the effect of all stimuli, which induce in human beings brightness- or darkness-excitations, upon the dark eye of the frog. The histological demonstration of the functional condition of the retina of the frog has, in the case of the investigations here reported, proved to be a fine test for the induction of brightness- or darkness-exciting stimuli.

The idea naturally presented itself, to inject the blood of light- or dark-adjusted human beings into the "dark" eye of the frogs. As expected, only in the first case did I find the frog's eye in a position of light-adjustment. Owing to external difficulties I have had to discontinue these experiments, and it has also been necessary to leave unexplained many other questions appertaining to the subject.

There is, for instance, the question of whether darkness-exciting stimuli effect only a diminution of brightness-excitation, or induce an active modification in the reverse direction. Also the question as to the existence of a "dark-hormone" has not yet been answered with certainty; the experiments which I have, myself, instituted in that connection speak *against* there being any such hormone. Experiments for the scrutiny of the retinas of frogs which have lived (in optical darkness) for several hours in a bright-smell or dark-smell atmosphere, I have also been unable to complete.

Conclusions

A variety of questions which have arisen in connection with the field of inquiry indicated in the present treatise have still to be cleared up. *The following has however been clearly ascertained.*

Brightness-exciting stimuli have an effect not only in the same manner on all sense-organs, not only from the sense-organs upon the muscles, and vice versa from the muscles upon the eye, but there is induced by such stimuli a modification of the *whole organism,*

which can be experimentally demonstrated on both human beings and animals by the most varied methods.

By *psychological* means we can establish brightness as a perception content; by *muscular-physiological* examination we can prove displacement of tone; by *histological* scrutiny of the retina of the frog we can ascertain contraction of the retinal cones and movement of pigment; on examination of the *blood* we have been able (hitherto by different means in the case of human beings and of frogs) to prove, by its effects, the presence of an as yet unknown kind of *hormone*. Brightness-excitation is at one and the same time a change of functional condition in the sense organs, in the bodily muscular system, in the animalic, and most probably also in the vegetative nervous system, as well as a hormonal change in the blood, characterized by the presence of a substance in all probability already latent in the bodily organs.

The question raised at the outset of this treatise, as to whether there exist "inter-modal" relations between the single domains of sense, must—as von Hornbostel has already shown—without doubt be answered in the affirmative. But in addition to that we have found that the capacity for reaction to brightness-stimuli is not only an essential characteristic of the sense-organs, as a *premodal function,* i.e., antecedent to all other sensory functions; but that brightness-excitation encompasses the *whole organism* in all its parts, as a fundamental biological process.[20]

References

1. Börnstein, W. Über die hormonale Steuerung der Korrelation zwischen den Sinnesleistungen und dem Tonus. (A lecture delivered at the Medizin.—Biologische Gesellschaft in Francfort-on-the-Main.) *Klin. Wochenschr.,* 1931, **28**.
2. ————. Über die funktionellen Beziehungen der Sinnesorgane unter einander und zum Gesamtorganismus. *Nederl. Tijdschr. v. Psychol.,* 1933, **1**.
3. Brentano, F. Untersuchungen zur Sinnespsychologie. Leipzig: 1907.
4. Dittler, R. In *Handb. d. norm. u. pathol. Physiol.,* ed. by Bethe, Bergmann, Embden, Ellinger. XII.
5. Englemann, T. W. *Arch. f. d. ges. Physiol.,* 1885, **35**.
6. Epstein, L. F. *Biol. Zsch.,* 1907.

[20]See also Börnstein, W. Der Aufbau der Functionen in der Hörsphäre. Berlin: Karger, 1930.

7. HELMHOLTZ, v. Die Tatsachen in der Wahrnehmung, 1878. (Ed. by Hertz and Schlick, Berlin, 1921.)
8. HORNBOSTEL, E. v. Über Geruchshelligkeit. *Pflüg. Arch. f. d. ges. Physiol.*, 1931, 227.
9. JAENSCH, E. R. Über den Aufbau der Wahrnehmungswelt. II. Leipzig: Barth, 1931. Pp. 594.
10. JAENSCH, W., & MANDOWSKY. Über die klinische Bedeutung psychischer Labilität bei optischen Wahrnehmungsvorgängen. Medizin. Welt. 32, 33.
11. KÖHLER, W. Akustische. Untersuchungen III. *Zsch. f. Psychol.*, 1915.
12. ————. Tonpsychologie. In *Handbuch der Neurologie des Ohres.* hrg. von Alexander, v. Marburg, Wien u. Berlin: Urban u. Schwarzinberg, 1923.
13. LASHLEY, K. S. Brain mechanism and intelligence. Chicago: Univ. Chicago Press, 1929. Pp. xi+186.
14. MACH, E. Die Analyse der Empfindungen. IV. Edit.
15. MANDOWSKY, C. Zur Klinik des Infantilismus. *Zsch. f. klin. Medizin*, **123**.
16. MERKULOW, J. Humorale Netzhautwirkungen. Versammlg. Deutsch. Natur forscher u. Arzte. Hamburg, 1928. *Klin. Wochen schr.*, **47**, 977.
17. METZGER, E. Experimentelle Untersuchungen über den Lichttonus des Manschen u. des Kaninchens. v. Graefe's *Arch. f. Ophthalmol*, 1931, **127**.
18. MICHAEL, S. Über die Beziehungen des Geschmacks zur chemischen Konstitution. *Biochem. Zsch.*, 1932, **255**.
19. MÜLLER, J. Zur vergleichenden Physiologie des Gesichtssinnes, etc., 1826.
20. PERKINS, F. T. A further study of configurational learning in the goldfish. *J. Exper. Psychol.*, 1931, **14**, 508-538.
21. QUADFASEL, Fr., & KRAYENBUHL, —. Über Haltungsstörungen u. ihre Beeinflussbarkeit. Eine klin. Untersuchung bei Kranken mit Torticollis spasticus mobilis. (Mon. *Zsch. f. Psychiat. u. Neurol.*, 1934. MSS.)
22. SCHILLER, P. v. Intersensorielle Transposition bei Fischen. *Zsch. f. vergl. Physiol.*, 1933, **19**.
23. ————. Das optische Verschmelzen in seiner Abhangigkeit von heteromodaler Reizung. *Zsch. f. Psychol.*, 1932, **125**, 249-289.
24. ————. Die Rauhigkeit als intermodale Erscheinung. *Zsch. f. Psychol.*, 1932, **127**.
25. STEIN, L. v. Der Schwindel. Leipzig and Moskau: 1908.
26. SZEKELY, L. Über den Aufbau der Sinnesfunktionen. *Zsch. f. Psychol.*, 1932, **127**.
27. ZIETZ, K. Gegenseitige Beeinflussung von Farb- und Tonerlebnissen. *Zsch. f. Psychol.*, 1931, **121**, 257-357.

London, England

THE FUSION FREQUENCY IN DIFFERENT AREAS OF THE VISUAL FIELD: I. THE FOVEAL FUSION FREQUENCY* [1]

From the Department of Psychology of Yale University

ROBERT T. ROSS

If the number of flashes per second emitted by a flickering test-patch be gradually increased, a frequency of intermittency may be found at which the test-patch loses its characteristics of flicker and appears steady or fused. Similarly, if the rate of intermittency be decreased from obvious fusion, a frequency may be found at which flicker becomes just perceptible. In general, these two critical frequencies are not the same; the first is usually higher (in flashes per second) than the second. If the fusion frequency is taken as the mean of these two threshold frequencies, the results are amenable to treatment by the Method of Limits, and it is possible to determine the fusion frequency with considerable accuracy.

There are two groups of conditions which might introduce variability into the determinations: (1) uncontrolled apparatus variables, and (2) variability and individual differences in subjects.

If variability is to be avoided, at least the following apparatus conditions must be controlled: brightness of the test-patch (47, 59, 48, 20, 13, 28, 30, 58, 50, 51, 10, 2, 3, 32),[2] color of the test-patch (54, 55, 52, 20, 53, 13, 30, 50, and 45-51 inclusive), size of the test-patch (26, 3), surrounds of the test-patch (44, 31), type of interruption of stimulus light (19, 22, 39, 14), speed of interruption of stimulus light (47, 48, 28, 29, 24), number of sectors in the interruption (7, 22, 9, 57, 28, 29, 8), and the relative brightness of the light and dark intervals (21, 58).

On the other hand it is necessary to consider those factors influ-

*Accepted for publication by Raymond Dodge of the Editorial Board and received in the Editorial Office, June 24, 1935.

[1]This study is from a doctoral dissertation presented to the Faculty of Yale University in May, 1934. The writer is indebted to Professor Walter R. Miles, under whose direction the investigation was pursued.

[2]References are given in chronological order of publication.

encing fusion frequency, the control of which is more a matter of methodology and choice of subjects than of apparatus construction. These conditions include: the state of adaptation of the eye (13, 36, 11, 1, 6, 41, 44, 27), the visual acuity (42, 62), the maintenance of fixation (57, 31), the pupillary area (28, 35, 60, 32), other stimuli in the visual field (7, 41, 25, 26, 34), pathological conditions of the optic mechanism (18, 19, 10, 43, 11, 44, 33), color-blindness (20, 49, 4, 42, 16), and the retinal region stimulated (12, 7, 17, 18, 19, 9, 49, 10, 5, 61, 11, 44, 15, 24, 32).

These lists of variables make no pretense of being complete. It is of the utmost importance, however, that both apparatus and subject variables be recognized and controlled. It is one of the misfortunes in optic investigations that independent studies are incomparable since the conditions of investigation have not been reproduced. Among the studies referred to, only those of comparatively recent date have made any attempt rigorously to control the experimental variables. Even in these contemporary studies, however, results have been accepted as valid when determined with only one or two subjects. As far as the writer is aware, the study by Lythgoe and Tansley (44) in which six subjects were used (one of whom proved anomalous) is the most complete survey on record of individual variability and difference in reporting fusion frequency.

Without implying any criticism of previous studies, it has seemed of interest to know something of the constancy of the fusion frequency. To that end, over 6500 individual readings of fusion frequencies, both at the fovea and in the peripheral retina, were made on 21 different subjects. From these data it has been possible to study individual variability and differences in the report of fusion frequency. The present section of this study deals with the variability of the foveal fusion frequency.

Choice of Subjects

In view of the long list of subjective factors which may influence the report of fusion frequency, it was thought best to limit the group of subjects, as far as possible, to "normal" persons with "normal" vision. Twenty-one university students in good health and doing creditable scholastic work were chosen. They ranged in age from 17 to 21 years, and had demonstrated their visual normalcy by

passing an ophthalmological examination according to the following standards: *acuity* 20/20 uncorrected, Snellen Chart; *astigmatism* not higher than first degree on Dodge astigmometer; *color vision* normal, Ishihara Tests for Color-blindness; *ocular dominance* determined by Miles V-Scope (A-B-C Test) (46) and dominant eye used in experiment; *fundus* normal, ophthalmoscopic examination; *color zones* normal, stereo-campimeter, Lloyd Charts and one-half degree test-patches of red, green, and blue; *muscular balance* normal, stereoscope and Wells Chart. To ten subjects selected at random the Bernreuter Personality Inventory and the Otis S-A Test of Mental Ability as adapted by Miles (45) were given.

The subjects formed a fairly homogeneous, visually, physiologically, and psychologically normal group of young adults. They cooperated well in the experiment and were paid for their services.

Apparatus and Method

As shown diagrammatically in Figure 1, the apparatus was housed in three rooms: a control room, an observation room, and a projection room. A shunt motor (*M*) (rated at 3000 r.p.m. at 110 volts)

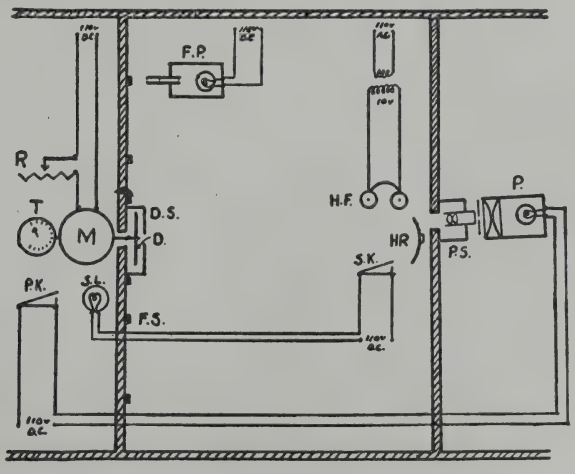


FIGURE 1
Diagrammatic View of Apparatus
See text for explanation

rotated a sectored disc (*D*) in the observation room. Securely fastened to the armature shaft was a tachometer (*T*) which gave motor speed in r.p.m. This tachometer could easily be checked by means of an electric revolution-counter and a stop watch. The speed of the motor was controlled by means of a rheostat (*R*) and a set of balanced resistances. With this arrangement the angular velocity of the disc could be varied at will and read at any instant. Conveniently located nearby were a single lamp (*SL*) operated by the subject by means of a key (*SK*), and a switch (*PK*) for turning the projection lantern (*P*) off or on.

The projection room housed only the projector (*P*) operated by the experimenter's key (*PK*). The projector carried a blackened metal slide pierced by a one-millimeter hole. The image of this hole was projected (through a two-centimeter square opening) into the observation room and focused on the sectored disc, where it presented a well-defined, circular test-patch 3.3 cm. in diameter. The projector was covered by a shield (*PS*) so that no stray light passed from the projection room into the observation room.

The observation room was fairly large, completely light-proof, and painted mat-black. Through the wall opposite the subject projected the shaft which carried the sectored disc (*D*). This disc was 25 cm. in diameter and consisted of two 180° sectors of mat cardboard, one white, with a reflection coefficient of 0.753, and one black, with a coefficient of 0.032. The entire surface of the disc was covered with a shield (*DS*), with the exception of a circular opening 12.5 cm. in diameter located directly above the center of the disc. This left the disc sufficiently exposed to project upon it the test-patch whose center was 6.25 cm. from, and immediately above, the disc's center. When the disc rotated slowly, the test-patch periodically appeared and disappeared as it was reflected now from the white, now from the black sector.

With the test-patch as their point of intersection, two lines, one vertical and one horizontal, were laid off on the walls of the room. Circular white discs one centimeter in diameter were attached to the wall at visual angles of 5, 10, 20, 30, and 45 degrees along the temporal and nasal coordinates of the retina and at 5, 10, and 20 degrees along the superior and inferior coordinates. These fixation spots (*FS*) could be illuminated singly by means of the projector

(*FP*) wihch gave a brightness of the spot less than 0.1 ml. but sufficient for foveal fixation.

The subject sat directly across the room in front of the test-patch. His head was steadied by the head-rest (*HR*) and the distance of his eye adjusted to 290 cm. from the test-patch. At this distance slight changes in the visual angle of the test-patch occasioned by a slight turning of the head with peripheral fixation were within the limits of error of the experiment.

With the eye at this distance the test-patch had a visual angle of 40 minutes. This size is sufficient to avoid the major difficulties encountered by small shifts in fixation when working with smaller patches, and to ensure visibility in both the fovea and periphery. On the other hand, it is small enough to allow measurement of the fusion frequency at five degrees in the periphery and to prevent apparent differences in fusion frequency at the center and edges of the patch.

The subject wore head-phones (*HF*) which ran directly off the 10-volt terminals of a transformer operated at 60 cycles and 110 volts. He was supplied with the signal key (*SK*) which operated the signal lamp (*SL*) and with a blindfold for his non-dominant eye.

The brightness of the test-patch, checked frequently, varied over 10 months from 2.42 to 2.48 with a mean value of 2.44 ml. The choice of this value was determined primarily by the projection lamp; the brightness could easily be kept constant, and could be checked with standard instruments. Moreover, the photon value with a seven-mm. pupil fell on the straight part of the curve which relates fusion frequency to photon illumination as determined by Hecht (32, p. 255). The room was illuminated during rest periods by white light giving an illumination at the eye of the subject of 3.42 ml. During the readings the room was dark and the subject saw only the fixation spot and the test-patch. The following experimental program was outlined for the subjects.

For two minutes the subject sits in the lighted room. During this time the blindfold is to remain off, as are the head-phones. At the end of the two-minute period the lights are turned off, the subject puts on the blindfold and head-phones, and adjusts himself to the head-rest. After one minute of darkness the test-patch, ob-

viously flickering, is lighted, and the frequency of its alternations is gradually increased until it fuses. At this point the subject pushes the signal key (*SK*) and the test-patch disappears [the experimenter having opened the projector key (*PK*) at the signal]. When next the patch appears it is fused. When first he becomes aware of flicker, the subject pushes the key, and the patch disappears. When next it appears it is flickering, and the cycle is repeated. At the end of five cycles (ten readings: five of fusion from flicker, five of flicker from fusion), the general illumination is turned on, headphones and blindfold removed for two minutes and the entire procedure repeated.

Each cycle of ten readings was completed in two minutes, a period which allowed six seconds for each reading and a six-second rest before the next. The time interval between the appearance of the test-patch and the subject's report was varied at random, but the rate of change of revolutions per second was practically constant for all readings. The variation in time-interval was accomplished by showing the test-patch at various degrees of remoteness from the threshold value and approaching the threshold value at a constant rate.

Treatment of Data. As explained above, ten readings with the test-patch passing from fusion to flicker were taken alternately with ten readings with the patch passing from flicker to fusion. As a result of this method, two thresholds were determined, (1) the fusion→flicker threshold and (2) the flicker→fusion threshold. Although the first was sometimes higher, sometimes lower, than the second, we can refer without undue confusion to the upper threshold and lower threshold as is usually the case with the method of limits. One set of ten readings thus determined an upper, the other set a lower, threshold. The most probable value of the fusion frequency was assumed to be the mean of these two values. The use of this measure needs no further justification other than to point out that if either set of readings were taken alone, the fusion frequency would appear either higher or lower than its most probable value.

A measure of threshold-difference (*D*) has been given with the mean values of the fusion frequency wherever this appeared desirable. This threshold-difference value is half of the total difference between the upper and lower threshold values whose mean determines a given fusion frequency. If, then, this value (*D*) be first added

and then subtracted from the fusion frequency, the actual values of the upper and lower thresholds may be obtained.

A statement of a measure of dispersion for the fusion frequency presented a more difficult problem. Standard deviations for the upper and lower thresholds may easily be found, but any combination of them to get a value of dispersion for their mean value is open to criticism. Inasmuch, however, as the values of D represent a constant error which, when added or substracted to the lower and upper thresholds respectively, transposes the means of the upper and lower threshold distributions to the value of the fusion frequency, it appears justifiable to consider the two threshold distributions as distributed about the fusion frequency value and to consider the measure of dispersion of the latter as identical with that of the combined upper and lower threshold distributions.

Now it may be shown that the average deviation of a distribution which consists in turn of two sub-distributions whose means are identical (i.e., $D=0$) is simply the mean of the A.D.'s of the two sub-distributions; while if D be so large that the two sub-distributions have no values in common, the average deviation of the entire group is numerically equal to D. If D is considered as a constant error, as it must be in this study, then, since after correction for the constant error D is always zero, the average deviation of the distribution will be simply the mean of the average deviations of the two sub-distributions.

In this work, therefore, the average deviation was calculated for the upper and lower thresholds, averaged, and converted to standard deviation (correction being made for the small number of cases).

The Variability of the Foveal Fusion Frequency

The nature of the data obtained by the Method of Limits makes it possible to discuss the problem of variability under three headings. (1) To what degree do 20 observations which determine a single value of the foveal fusion frequency for a given subject at one sitting agree among themselves? (2) To what degree does the foveal fusion frequency vary in the same subject from session to session? (3) To what degree does the foveal fusion frequency vary from subject to subject?

The fourth column of Table 1 gives the standard deviation of readings about their mean values. It is to be remembered that these

TABLE 1

THE MEAN FOVEAL FUSION FREQUENCIES OF TWENTY-ONE NORMAL SUBJECTS

Case no.	No. of readings	F.F.F. (flashes per second)	Mean σ	Mean D
1	40	30.05	0.607	0.33
2	220	29.18	0.485	0.57
3	230	24.78	0.488	0.32
4	240	28.32	1.033	0.89
5	220	25.42	1.057	1.58
6	280	26.90	0.743	0.73
7	160	26.15	0.850	0.42
8	160	28.73	0.570	0.27
9	40	27.10	0.977	0.17
10	200	26.38	0.645	0.33
11	200	26.20	0.758	0.55
12	120	24.35	0.833	0.28
13	120	26.27	0.827	0.82
14	80	27.18	0.702	1.75
15	40	26.78	0.760	0.18
16	40	23.49	0.598	1.02
17	40	23.55	1.182	0.23
18	40	27.12	0.672	0.41
19	40	22.79	0.725	0.27
20	40	25.58	0.991	0.15
21	40	24.70	0.790	0.46
Mean	N = 2590	26.13	0.725	0.56
		$\sigma_{dis.} = 1.836$		
		$\sigma_M = 0.405$		

readings are considered in groups of ten and the sigmas (subject to student's correction) for the group averaged. In Case 1, for instance, four sigmas were calculated, one for each threshold determination, and the four averaged. This gives the most probable value of the actual variability. Column five gives the average amount by which the mean foveal fusion frequency for any group of ten differed from the mean for all readings as given in column three. We may answer the first question, then, as follows: On the basis of 2590 readings, made on 21 subjects, we conclude that any one threshold determination (i.e., ten readings, flicker to fusion or ten readings, fusion to flicker) has a standard deviation of 0.725, and that any such threshold value differs, on the average, from the foveal fusion frequency as determined by two such thresholds by 0.56 flashes per second.

TABLE 2

THE FOVEAL FUSION FREQUENCY AND ITS VARIABILITY FOR INDIVIDUAL CASES

Case	No. of series	F.F.F. (flashes per sec.)	σ (corrected)
2	11	29.18	0.663
3	11	24.78	0.703
4	12	28.32	2.349
5	10	25.42	1.082
6	14	26.90	1.469
10	10	26.38	1.393
11	10	26.20	0.424
Mean		26.77	1.155

The results given in Table 2 cast some light on the second question. For Case 2, for example, 11 determinations of the foveal fusion frequency (each determination based, in turn, on 20 individual readings) had a standard deviation of 0.663. The most probable value of the standard deviation of the foveal fusion frequency for any individual is 1.155. The chances are even, therefore, that any single determination of foveal fusion frequency may differ from the most probable value by as much as 0.779 flashes per second.

The answer to the third question may be found by again referring to Table 1, column three. Each value of the foveal fusion frequency given is the most probable value for the given case on the basis of the readings made. For 21 cases the standard deviation of the distribution is 1.836 flashes per second, and the standard error of the mean, 0.405. This indicates that deviations as large as 6.6 x 1.836 or 12.118 flashes per second between individuals may be attributable merely to individual differences.

These results are summarized in Figure 2. The data relevant to questions two and three are explicit in this figure. It should be kept in mind that each block in Figure 2 represents the mean of 20 readings, ten of which measured the flicker-fusion threshold and ten measured the fusion-flicker threshold (with the sigmas indicated in Table 1, column four), and that the mean values of these ten readings deviated from the mean value of the 20 by an amount indicated in Table 1, column five.

In general, therefore, we expect a relatively high degree of constancy in report of one subject at one sitting ($\sigma = 0.725$); we expect

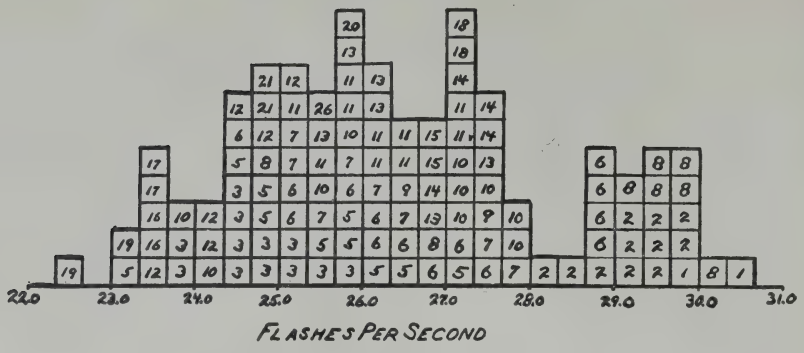


FIGURE 2

THE DISTRIBUTION OF INDIVIDUAL VALUES OF THE FOVEAL FUSION FREQUENCY
Numbers in the squares refer to case number. Each square represents the mean of 20 readings taken at one session.

less agreement between foveal fusion frequencies determined on the same subject at different sittings ($\sigma = 1.155$), and still less agreement between the mean foveal fusion frequencies of different subjects ($\sigma = 1.836$).

We may conclude that, in any satisfactory attempt to study flicker phenomena, the most consistent results will be obtained by performing the experiments on a given individual at a single session. The findings also seem to indicate that the constants which must be applied to any theoretical formulation of visual function must be determined for each individual and for the same individual at different times.

We are thus brought to regard the visual process in its physiological aspects. Burch (13) maintained that variability in judgment and change in fusion frequency with physiological status were both so considerable as practically to invalidate any attempt at measurement. Ferry (20) and Porter (51), on the other hand, maintained that fusion frequencies were practically independent of the individual's condition. Wiersma (61) has shown that fusion frequency is different for different types of psychopathic condition, but whether the difference is in judgment or in physiological condition is, of course, unknown.

We should expect, of course, to find changes in fusion frequency in any pathological condition which affects the visual mechanism.

Exner (18, 19) reported decreased fusion frequency with glaucoma. Braunstein (10, 11) found, in part, that atrophy of the optic nerve and certain types of central scotoma and pigment degeneration in the retina decrease the fusion frequency. Helson and Guilford (33) confirmed this latter finding by demonstrating that the fusion frequency is higher in pigmented eyes. Lythgoe and Tansley (44) have demonstrated anomalous behavior with night-blindness. In view of these findings and the many physiological conditions which may affect the visual apparatus—especially changes in blood-pressure with attendant changes in the concentration and consistency of the ocular fluids—it becomes increasingly apparent that some method of demonstrating the "normalcy" of any given vision is demanded. In any event, it is well to ask for some assurance of the normalcy of subjects in regard to their visual function before results are accepted.

Summary

The foveal fusion frequencies of 21 normal subjects were determined. The test conditions were constant throughout and involved the use of a test-patch 40 minutes of visual angle in diameter having a brightness of 2.44 ml. and viewed against dark surrounds after two minutes of adaptation to an illumination of 3.42 ml. and one minute in the dark.

1. Flicker-fusion threshold and fusion-flicker threshold readings had an average standard deviation of 0.725 flashes per second.

2. Flicker-fusion threshold and fusion-flicker threshold differed from the fusion frequency as found by averaging the two thresholds by 0.56 flashes per second on the average.

3. A series of determinations of the fusion frequency (20 readings per determination) for a given subject showed an average dispersion (S.D.) of 1.155 flashes per second over a period of months.

4. The mean fusion frequencies for 21 subjects showed a dispersion (S.D.) of 1.836 flashes per second.

It is concluded that psychological and physiological differences in individual subjects make comparison of absolute values of fusion frequency inadvisable. Results obtained from a sufficient number of readings on a single subject at a single session are probably amenable to theoretical treatment, provided the general physiological and psychological state of the subject does not materially change nor indicate pathology.

References

1. Adams, D. R. Dark adaptation (a review of the literature). *Med. Res. Council, Great Britain, Spec. Rep. Ser.,* 1929, No. 127. Pp. 138.
2. Adrian, E. D., & Matthews, R. The action of light on the eye. I. The discharge of impulses on the optic nerve and its relation to the electrical charges in the retina. *J. Physiol.,* 1927, **63**, 378-414.
3. ————. The action of light on the eye. II. The process involved in retinal excitation. *J. Physiol.,* 1927, **64**, 279-301.
4. Allen, F. Effect upon the persistence of vision of exposing the eye to light of various wave lengths. *Phys. Rev.,* 1900, **11**, 257-290.
5. ————. Some phenomena of persistence of vision. *Phys. Rev.,* 1909, **28**, 45-56.
6. ————. On reflex visual sensations. *J. Opt. Soc. Amer.,* 1923, **7**, 583-626.
7. Aubert, A. Physiologie der Netzhaut. Breslau: Morgenstern, 1865. Pp. 394.
8. Basler, A. Ueber die Verschmelzung von zwei nacheinander erfolgenden Lichtreizen. *Pflüg. Arch. f. d. ges. Physiol.,* 1912, **143**, 245-251.
9. Bellarminow, L. Ueber intermittierende Netzhautreizung. *Graefes Arch. f. Ophth.,* 1889, **35**, 25-49.
10. Braunstein, E. F. Beitrag zur Lehre des intermittirende Lichtreizes der gesunden und kranken Retina. *Zsch. f. Psychol.,* 1903, **33**, 171-206, 241-288.
11. ————. Zur Lehre von den kurzdaurenden Lichtreizen der Netzhaut. *Zsch. f. Sinnesphysiol.,* 1923, **55**, 185-229.
12. Brewster, D. On the duration of luminous impressions on certain points of the retina. *Brit. Asso. Rep.,* 1858, **28**, Pt. II, 6-10.
13. Burch, G. J. An account of certain phenomena of color vision with intermittent light. *J. Physiol.,* 1897, **21**, 426-434.
14. Cobb, P. W. Some comments upon the Ives theory of flicker. *J. Opt. Soc. Amer.,* 1934, **24**, 91-98.
15. Creed, R. S., & Ruch, T. C. Regional variations in sensitivity to flicker. *J. Physiol.,* 1932, **74**, 407-423.
16. Dow, J. S. Some further notes on the physiological principles underlying the flicker photometer. *Phil. Mag.,* 1910, **19**, 58-77.
17. Exner, S. Bemerkungen ueber intermittirende Netzhautreizung. *Pflüg. Arch. f. d. ges. Physiol.,* 1870, **3**, 214-240.
18. ————. Zur Kentniss von der Regeneration in der Netzhaut. *Pflüg. Arch. f. d. ges. Physiol.,* 1879, **66**, 407-409.
19. ————. Weitere Untersuchen ueber die Regeneration in der Netzhaut und ueber Druckblindheit. *Pflüg. Arch. f. d. ges. Physiol.,* 1879, **20**, 614-626.
20. Ferry, E. S. Persistence of vision. *Amer. J. Sci.,* 1892, **44**, 192-207.
21. Flick, A. E. Ueber den zeitlichen der Erregung in der Netzhaut. *Arch. f. Anat., Physiol., u. Wiss. Med.,* 1863, 739-764.

22. FILEHNE, W. Ueber den Entstehung des Lichtstaubes, der Starrblindheit und der Nachbilder. *Graefes Arch. f. Ophth.,* 1885, **31**, Abt. II, 1-30.

23. FRY, G. A. Color phenomena from adjacent retinal areas for different temporal patterns of intermittent white light. *Amer. J. Psychol.,* 1933, **45**, 714-721.

24. GELDARD, F. A. Foveal sensitivity as influenced by peripheral stimulation. *J. Gen. Psychol.,* 1932, **7**, 185-189.

25. GRANIT, R. Comparative studies on the peripheral and central retina. I. On the interaction between distant areas in the human eye. *Amer. J. Physiol.,* 1930, **94**, 41-50.

26. GRANIT, R., & HARPER, P. Comparative studies on the peripheral and central retina. II. Synaptic reactions in the eye. *Amer. J. Physiol.,* 1930, **95**, 211-228.

27. GRANIT, R., & HAMMOND, E. L. Comparative studies of the peripheral and central retina. V. Time course of fusion frequency. *Amer. J. Physiol.,* 1931, **98**, 654-663.

28. GRUENBAUM, O. F. F. On the intermittent stimulation of the retina, I. *J. Physiol.,* 1897, **21**, 396-402.

29. ————. On the intermittent stimulation of the retina, II. *J. Physiol.,* 1897, **22**, 433-450.

30. HAYCRAFT, J. B. Luminosity and photometry. *J. Physiol.,* 1897, **21**, 126-146.

31. HECHT, S., SHLAER, S., & VERRIJP, C. Intermittent stimulation by light. II. The measurement of critical fusion frequency for the human eye. *J. Gen. Physiol.,* 1933, **17**, 237-249.

32. HECHT, S., & VERRIJP, C. Intermittent stimulation by light. IV. A theoretical interpretation of the quantitative data of flicker. *J. Gen. Physiol.,* 1933, **17**, 269-282.

33. HELSON, H., & GUILFORD, J. P. The relation of visual sensitivity to the amount of retinal pigmentation. *J. Gen. Psychol.,* 1933, **9**, 58-76.

34. IVES, H. E. Studies in the photometry of light of different colors. I. Spectral luminosity curves obtained by the equality of brightness photometer and the flicker photometer under similar conditions. *Phil. Mag.,* 1912, **24**, 149-188.

35. ————. Studies in the photometry of lights of different colors. II. Spectral luminosity curves by the method of critical frequency. *Phil. Mag.,* 1912, **24**, 352-370.

36. ————. Studies in the photometry of lights of different colors. III. Distortions in spectral luminosity curves produced by variations in the character of the comparison standard and of the surroundings of the photometric field. *Phil. Mag.,* 1912, **24**, 744-751.

37. ————. Studies in the photometry of lights of different colors. IV. The addition of luminosities of different color. *Phil. Mag.,* 1912, **24**, 845-853.

38. ————. Studies in the photometry of lights of different colors. V. The spectral luminosity curve of the average eye. *Phil. Mag.,* 1912, **24**, 853-863.

39. ————. Critical frequency relations in scotopic vision. *J. Opt. Soc. Amer.,* 1922, **6**, 254-268.

40. ————. A theory of intermittent vision. *J. Opt. Soc. Amer.,* 1922, **6**, 343-361.

41. KLEITMAN, N., & PIÉRON, H. De la variation du taux de sommation superficielle des impressions lumineuses en fonction de la nature de la lumière, de la region retinienne excités et de l'état d'adaptation. *C. r. Soc. de biol.,* 1929, **100**, 1174-1177.

42. VON KRIES, J. Ueber die Wahrnehmung des Flimmers durch normale und durch farbenblinde Personen. *Zsch. f. Psychol.,* 1903, **32**, 113-117.

43. LOHMANN, W. Ueber die lokalen Unterschiede der Verschmelzungsfrequenz auf der Retina und ihr abweichendes Verhalten bei der Amblyopia congenita. *Graefes Arch. f. Ophth.,* 1908, **69**, 395-411.

44. LYTHGOE, R. J., & TANSLEY, K. The adaptation of the eye: its relation to critical frequency of flicker. *Med. Res. Council, Great Britain, Spec. Rep. Ser.,* 1929, No. 134. Pp. 72.

45. MILES, C. C. The Otis S-A as a fifteen-minute intelligence test. *Person. J.,* 1931, **10**, 246-249.

46. MILES, W. R. Ocular dominance demonstrated by unconscious sighting. *J. Exper. Psychol.,* 1929, **12**, 113-126.

47. PLATEAU, J. Ueber einige Eigenschaften der vom Lichte auf das Gesichtsorgan hervorgebrachten Eindruecke. *Ann. d. Physik. u. Chem.,* 1830, **20**, 304-332. (Translation of "Dissertation sur quelques propriétes des impressions produites par la lumière sur l'organe de la vue," Liège, 1829.)

48. ————. Betrachtungen ueber ein von Hrn. Talbot vorgeschlagenes photometrisches Princip. *Ann. d. Physik u. Chem.,* 1835, **35**, 457-468.

49. POLIMANTI, O. Ueber die sogenannte Flimmer-Photometrie. *Zsch. f. Psychol.,* 1899, **19**, 263-283.

50. PORTER, T. C. Contributions to the study of flicker, I. *Proc. Roy. Soc. London,* 1898, **63**, 347-356.

51. ————. Contributions to the study of flicker, II. *Proc. Roy. Soc. London,* 1902, **70**, 313-329.

52. ROOD, O. N. Modern chromatics. New York: Appleton, 1879. Pp. 329.

53. ————. On a photometric method which is independent of color. *Amer. J. Sci.,* 1893, **46**, 173-176.

54. SCHAFHAUTL, ——. Abbildung und Beschreibung des Universals—Vibrations—Photometers. *Acad. Wissen. München,* 1854, **7**, Abt. II, 465-497.

55. ————. Über Photometrie. *Acad. Wissen. München,* 1854, **7**, Abt. II, 499-525.

56. SCHATERNIKOFF, M. Über den Einfluss der Adaptation auf die Erscheinung des Flimmers. *Zsch. f. Psychol.,* 1902, **29**, 241-254.

57. SCHENK, F. Über Intermittirende Netzhautreizung. I. Ueber den Einfluss von Augenbewegungen auf die Beobachtung rotirender Scheiben zur intermittirenden Netzhautreizung. *Pflüg. Arch. f. d. ges. Physiol.,* 1896, **64**, 165-178.

58. SHERRINGTON, C. S. On reciprocal action in the retina as studied by means of some rotating discs. *J. Physiol.*, 1897, **21**, 33-54.

59. TALBOT, H. F. Experiments on light. *Phil. Mag.*, 1834, **5**, 321-324.

60. TROLAND, L. T. The theory and practice of the artificial pupil. *Psychol. Rev.*, 1913, **22**, 161-176.

61. WIERSMA, E. D. Lectures on psychiatry. London: Lewis, 1932. Pp. 601.

62. WOOG, P. De la persistance variable des impressions lumineuses sur les differentes regions de la retine. *C. r. Acad. sci.*, 1919, **168**, 1222-1224.

Yale University
New Haven, Connecticut

THE TOUCH SPOTS AND THE INTENSITY OF THE STIMULUS*

From the Psychological Laboratory, University of Nebraska, Lincoln, Nebraska

J. P. Guilford and Edna M. Lovewell

Many a student in the elementary laboratory course has been called upon to make a fine cross-section map on the skin of his *O*'s forearm, to explore that region point by point with a tactile or temperature stimulus in the search for the well-known sensitive "spots" therein. The same student is usually given the implicit, if not explicit, assurance that underneath each such sensitive "spot" lies a single receptor cell. Titchener, for example, tells the student, "This experiment enables us to discover, and accurately record, the sense-organs of temperature" (8, p. 54). And again, in speaking of the touch spots, "This experiment answers the same questions in regard to the sense of pressure . . ." (8, p. 58). And Meyers tells the student, "The skin contains certain 'spots' which are particularly sensitive to the lightest touch, others which are sensitive to pain, others again are sensitive to cold and heat" (5, p. 110). Foster and Tinker are even more explicit when they say, "Even the four cutaneous sensations mentioned cannot be evoked uniformly over the whole surface of the body, but appear *only in certain very small areas,* or 'spots' arranged irregularly *with far larger areas of insensitivity between*"[1] (2, p. 105). Many authors follow Titchener in giving the approximate number of "spots" per sq. cm. for the different kinds of sensory qualities (9, p. 147 ff.). The numbers given are touch spots, 25; cold spots, 13; warm spots, 2; and pain spots, 100 to 200. These are averages, of course, since different regions of the body are unequally sensitive. On the whole there is no general agreement in the literature as to the average number of "spots" for any specified tactual area.

It is this last conception in particular, that the number of sensi-

*Accepted for publication by John Paul Nafe of the Editorial Board and received in the Editorial Office, June 11, 1935.

[1]The italics are ours.

tive spots can be determined for a specified area, that we wish to question in this study. We wish also to point out other aspects of these traditional notions that seem to us to be erroneous. That the skin is not uniformly sensitive for any given stimulus, even within the same sq. cm., we are quite willing to admit. That there are definite numbers of restricted, isolated "spots" that can be located and counted by means of our present methods of stimulation we seriously doubt. A number of investigators, particularly Dallenbach (1), have shown conclusively that for each cold and warm "spot" there is not a single specialized receptor. Furthermore, with meticulously controlled conditions, Dallenbach was unable to obtain the traditional numbers, relatively, of warm and cold spots. On the upper arms of three *O's*, he found on the average 12, 14, and 18 cold spots, which is not far from the average generally accepted. But he found also averages of 7.5, 22.5, and 20.0 warm spots respectively in the same patches of skin (1, p. 416). The last figures are significantly remote from the traditional two warm spots per sq. cm. Nafe, after much consideration of the problem, comes to the conclusion that "The idea . . . that a punctiform stimulation arouses impulses over a single nerve (*sic*) is absurd" (6, p. 388). It may be that the points which we wish to urge in connection with cutaneous sensitivity are already generally recognized. But even recently Wheeler (10, p. 438) takes great pains to explain why cutaneous "spots" are round. "The roundness," he says, "illustrates the condition of least action . . . When a single nerve ending or a group or nerve endings is stimulated, . . . the resulting pattern of nervous energy adjusts itself to the patterns of equal potential around it. . . . If this is true, we might expect the "spots" to be round." This seems to the writers to be a rather unsatisfactory attempt to explain something that needs no explaining, since the traditional "spots" do not exist in the form of "polka dots" as Wheeler seems to assume. It may be that Wheeler is confusing the physiological "spots" with the felt experience, which may indeed be round for round pointed stimulus.

The problem arose for the authors in the elementary laboratory. In the first place, students have great difficulty in deciding how many touch or temperature "spots" they have found in a mapped portion of the skin. Suppose that they have stimulated every sq.

mm. just once. Is each responsive sq. mm. to be regarded as *one* sensitive spot? Suppose, on the other hand, that they have stimulated each sq. mm. four times. Does each quarter of a sq. mm. which gives a response deserve to be called a "spot"? When each sq. mm. is stimulated only once, there is sometimes a tendency for the responsive squares to be clustered. Should each cluster be called a "spot"? If this rule is adopted there are still difficulties. The clusters have very irregular shapes and also a very disturbing habit of merging into one another at places; they are by no means completely segregated. What then becomes of the hypothetical isolated sensitive "spots"? And how can one succeed in counting such "spots"?

Another common question arising concerns the intensity of the stimulus that should be used. Experimental method and common sense demand a uniform stimulus. But how strong should the stimulus be? Laboratory manuals are not very specific as to this phase of the experiment. Titchener, in speaking of tactile stimuli, makes no recommendation except that "The point of the instrument must make a noticeable depression in the skin" (8, p. 59). This surely leaves to *E* a great latitude. Foster and Tinker are a little more specific when they say, "Try a horse hair projecting 1¼ inches" (2, p. 109). They add the suggestion that if this is always felt one should lengthen the hair; how much, they do not say. The implicit assumption seems to be that the touch "spots" may be located and counted if one uses a stimulus that is neither too weak never to be felt nor too strong always to be felt. Some of our overly curious students have frequently discovered that the number of spots that one seems to find will depend upon the strength of the stimulus. This has led us to wonder whether there is a definite quantitative functional relationship between these two variables, and to call into question the whole traditional conception of the sensitive "spots." Why could not one regard the sensitivity of the skin from a statistical point of view, defining sensitivity of each unit area as the probability of a response with a given intensity of stimulus? One would then expect that the functional relationship between the responsive proportion of the total area and the intensity of the stimulus would be of the ogive type, conforming to the phi gamma hypothesis. This is essentially what we have found to be true. And from this fact, and from others, certain suggestions can be made concerning theories of the cutaneous sense or senses.

Procedure

Stimuli. Horse hairs of uniform quality and diameter, as attested by microscopic examination and measurement, were "weighed" on sensitive balances until the proper lengths (intensities) were found. The hairs were then fastened to the ends of match sticks with sealing wax. The intensities approximated a geometric series, being of the following magnitudes: .01, .02, .04, .10, .20, .40, .60, 1.00, and 1.60 grams.

Observers. Two of the three *O*'s were unaccustomed to any type of psychological observation but adjusted themselves readily to the task. The third *O* was an instructor in psychology who had had much experience in the laboratory but who had not observed to any great extent in tactual problems. Every *O* came to the laboratory at the same time each day, five days of the week, during July of the summer session of 1932. The centimeter square patch that was being investigated with each *O* was explored completely with every stimulus-intensity at least twice. The weaker stimuli were applied to the whole area on as many as five different days. The area was completely explored with each of from two to four different intensities on the same day. An attempt was made to combine different intensities on different days.

Method. *O* would lay his hand, slightly clenched, dorsal side up, on the table. A selected area on the back of the hand was shaved free of hair. A map was applied with the usual rubber stamp with sq. mm. divisions. *E* selected the stimulus to be used and after a "Ready" signal began the exploration. The stimulus was applied perpendicularly and with just enough force to make a perceptible bend in the hair. A magnifying glass aided in the accurate placing of the end of the hair. The hair was applied for about one second at each intersection of the lines of the map and also at the centers of the squares. It was thus applied 200 times within the area. When *O* reported "There," *E* stopped to record the response. If there was an after-sensation, *O* reported "Persists" and there was a pause to allow time for recovery. This pause was required very infrequently while using the weaker stimuli but very often with the stronger ones. No extraordinary care was taken to place the map on the identical points on successive days since we

were not interested particularly in the verification of the *location* of sensitive "spots." Since *O* came to the experiment almost every day, however, it was an easy matter to find traces of preceding maps at each new sitting.

Results. The quantitative results are given in Table 1 and in Figure 1. In both cases we have given the proportions of the 200 stimulations which gave tactual responses, for every stimulus-intensity and for every *O*. At this point we shall have to confess that the responses were not purely "tactual" in the strictest sense. The sensory responses were, for the stronger stimuli, at least, often contaminated with pain. No pain responses were given for stimuli weaker than .04 gram, and such responses were rare up to an intensity of .40. When reports of pain first made their appearance, we asked the *O*'s to distinguish between the two kinds of report. They soon found this impossible in many cases. Very often there would be mixtures of qualities. Then, frequently it happened that, if no pain was felt at the moment of stimulation, it would crop up later and continue as an after-effect. Such a mingling of pressures and pains, and the fact that stronger stimuli gave a higher probability of pain, point toward the quantitative theory of "felt experience" as proposed by Nafe (6, 367 ff.). All felt experiences, according to Nafe, are "patterns of brightness (or sharpness) of greater or less complexity" (6, p. 389), analogous to complex tones with their different timbres. The effect of intensifying a stimulus is to increase the brightness or sharpness of the experience. This change has commonly been described as a shift from pressure to pain. If Nafe is correct, we are perfectly justified in pooling all our reports of experience, either of pressure or of pain, as, in fact, we had to do. If there are two distinct receptors for pressure and pain, however, we are not justified in doing so. We believe that the continuity of the curves in Figure 1 will lend support to the conclusion that we are dealing with one quantitative variable and not two. If we were actually dealing with two separate senses, then we would have in our composite function two ogives of differing median and slope, and the combined result would have in it a "plateau" or at least an interrupted rise near its center. No such interruption is evident in the curves of Figure 1.

Both Table 1 and Figure 1 show strikingly the continuous in-

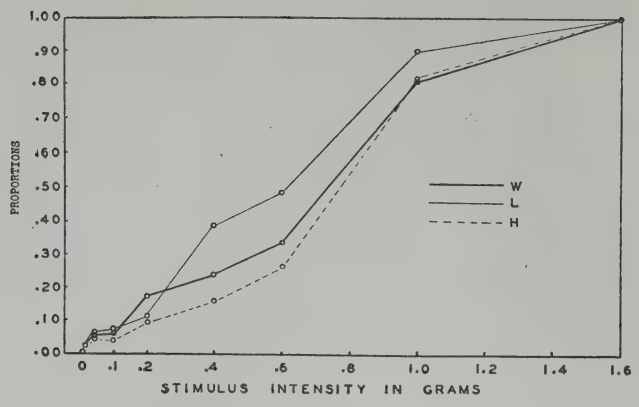


FIGURE 1

INCREASE IN THE PROPORTION OF RESPONSES WITH INCREASE IN THE INTENSITY OF THE TACTUAL STIMULUS

TABLE 1

PROPORTIONS OF ALL THE STIMULATIONS WITH EVERY STIMULUS-INTENSITY THAT GAVE RESPONSES IN THE SAME AREA ON THE SKIN

	Stimulus-intensity in grams								
	.01	.02	.04	.10	.20	.40	.60	1.00	1.60
L	.010	.025	.066	.075	.120	.388	.488	.907	1.000
H	.009	.025	.052	.042	.095	.163	.267	.823	1.000
W	.003	.025	.045	.055	.173	.240	.338	.818	1.000
Mean	.007	.025	.053	.057	.129	.264	.367	.839	1.000

crease in the number of responses (spots?) as one increases the stimulus-intensity. The curves in Figure 1 are all very similar and they are apparently of the ogive type.

Discussion. How shall one count the number of "tactile" "spots" in any one of the three areas explored? What strength of stimulus should one adopt in determining the number? If one adopts the liminal value, 50 per cent of the area would respond. This would give the same number of "spots" for every kind of stimulus, provided we count every responsive point as one "spot." When one attempts to group the responses, as we have said before, one fails again because one cannot tell where one spot leaves off and another begins. As we study the successive records of the same area, going

from the weakest to the strongest stimulus, we do not find that the responding points which come in later for the stronger stimuli merely cluster around the scattered ones already present for the weaker stimuli. The spread in sensitivity is not merely from certain centers or "foci" of greatest sensitiveness as the traditional conception would lead us to expect. The new responding points crop up in quite unexpected places, in the midst of formerly insensitive patches. Former clusters break up and new ones form as the same area is explored with different stimuli. This happened in our own experiments on the same day so there is no question of the map's shifting to a new position on the skin.

It seems to the writers that the most reasonable and conservative conclusion to make is that every point on the skin has a given probability of responding to each stimulus-intensity. A point on the skin that is repeatedly stimulated does not obey an all-or-none law, either always responding or else never responding, for a superliminal stimulus, as the traditional "spot" hypothesis would seem to demand. At one moment the point gives a response and at another moment it does not. We would like to propose this fact as an explanation of the imperfect correlation that is usually found in the location of the "spots" in successive mappings. Even Dallenbach, who has probably more carefully duplicated conditions in successive mappings than any one else, found the agreements between two maps, when compared point for point for warm and cold spots, to vary from 80 to 96 per cent (1, p. 422). Other investigators have found far less correlation. And when Dallenbach compares four successive mappings, the percentage of identity drops down to a range of 73 to 88 (1, p. 422). In a larger number of successive maps the percentage would undoubtedly fall to a lower figure. All this, we maintain, is at least partly due to the fact that every stimulated point has its own threshold and its own psychometric function. It is conceivable that if the uncertainties of the epithelial tissues which lie above the receptors and through which the stimulus must be exerted could be ruled out, the percentage of identity in successive maps could be raised to 100. But we doubt whether even this would be true.

Is it not likely that the receptors of the skin are similar to others, to those of the retina, for example, as regards thresholds? Hecht

(3, pp. 253-262) presents convincing evidence to show that each visual receptor has its own threshold value and that the receptors taken as a group would form a normal distribution when ranged according to sensitiveness. An increased stimulus therefore brings into play more and more receptors as their threshold values are exceeded. Each receptor also responds with a higher frequency when increased energy is imposed upon it. Thus visual brightness is increased in two ways, by more receptors' responding and by an increased frequency in each fiber which responds.

Nafe (7) accepts these recognized properties of nervous responses as applying to cutaneous receptors. "Stimulation, even of comparatively small areas," says Nafe, "is usually accompanied by the initiation of impulses over a large number of fibers" (7, p. 393). Even a pointed stimulus affects more than one receptor, due to the fact that it sets up a mechanical gradient on the skin, with radiating stresses in the surrounding tissues. And there is little doubt that intensifying the stimulus increases the range of influence of this gradient, bringing more receptors within reach of the stimulus. This is one additional method, not applicable to the retina, of increasing the possibility of exciting a larger number of receptors and so of raising the probability of a response. But we maintain that this mechanical spread alone is not sufficient to account for the facts represented in Figure 1. Some additional observations of our *O*'s bear upon this point. We would alternate in succession the strongest stimulus, 1.6 gm., and one of the weakest, .04. The *O*'s general impression was that the area of the former was about three times that of the latter. Let us assume that this represents a fair indication of the spread of the mechanical gradient. A much better measure of this spread would be, of course, some physical measurement of the gradient itself. For these same two stimuli the probabilities of a response were 1.00 and .053, respectively. The former stimulus gave about 20 times as many responses as the latter, although the felt area was only about three times. It is for this reason that we say that the spread of the mechanical gradient is not sufficient in itself to account for the increase in the probability of a response. The latter must also be due in part to the other two factors, namely, the arousing of more receptors as their thresholds are exceeded and the increase in frequency of impulses in the same

nerve fibers. Work should be done on the effectiveness of the spread of the mechanical gradient as more pressure is applied to the stimulus.

In place of the time-honored conception of sensitive "spots" we would like to substitute a more accurate notion of "gradients of sensitivity" (not to be confused with the mechanical gradients discussed above). That the skin is not homogeneously sensitive, no one can deny. That there may be relatively isolated foci of greatest sensitivity, with lesser sensitivity between, we may have to admit, in view of Dallenbach's careful verification of such clusters of sensitive points for warm and cold. This arrangement we may be properly justified in calling "gradients of sensitivity." In order to secure a much better idea of the distribution of sensitivity, one should perform the tedious experiment of measuring the threshold at the center of every sq. mm. in a well-mapped portion of the skin. One could then draw a topographical map of the area. We would predict that if there are "peaks" of higher sensitivity these peaks are of different "altitude." How much of this topography is due to the arrangement and the properties of the receptor cells beneath the skin and how much of it is due to the rather indirect methods of stimulation, is still another problem. To settle this question one would need to strip the overlying tissues away and to repeat the experiment upon the bared receptors, if this were possible. In the meantime it is probably safe to postulate a topography of differential sensitivity for the receptors themselves that is consistent with the known facts of nervous conduction. In the light of these considerations, we maintain that the time-honored conception of sensitive "spots" be severely revised.

Summary and Conclusions

Patches of cutaneous surface one centimeter square were explored repeatedly with stimuli varying in intensity from .01 to 1.60 gm. Of the 200 points stimulated within these areas, the number or proportion giving a response increases as the stimulus-intensity increases. The proportion is less than one per cent for the weakest stimulus and 100 per cent for the strongest. The function is a continuous one of the ogive type. The following conclusions may be drawn.

1. The traditional cutaneous "spots" are largely fictions. We may be justified in speaking of "gradients of sensitivity" but not of "spots."

2. Such "gradients of sensitivity" are probably a property of the the distribution of sensitivity of the cutaneous receptors, some receptors having lower thresholds than others.

3. Under the present conditions of cutaneous stimulation, indirect as it is, we are strictly justified only in speaking of the probability of a response for a given point on the skin and for a given stimulus-intensity. The "gradients of sensitivity" as we know them empirically are gradients of probability.

4. An increased intensity of the stimulus probably has three effects, namely, (1) additional receptors with higher thresholds are brought into play; (2) the frequency of impulses in receptors already acting is increased; (3) due to the mechanical gradient set up by the stimulus, a wider area of tissue is affected and more remote receptors are brought into play.[2]

References

1. Dallenbach, K. M. The temperature spots and the end-organs. *Amer. J. Psychol.,* 1927, **39**, 402-427.
2. Foster, W. S., & Tinker, M. A. Experiments in psychology. (Rev. ed.). New York: Holt, 1929. Pp. xvi+392.
3. Hecht, S. The nature of the photoreceptor process. In *The foundations of experimental psychology,* ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1929. Pp. 216-272.
4. Heiser, F. Stimulus temperature and thermal sensation. *Arch. Psychol.,* 1932, No. 138. Pp. 81.
5. Myers, C. S. A textbook of experimental psychology. New York: Longmans, Green, 1925. Pp. 358.
6. Nafe, J. P. The psychology of felt experience. *Amer. J. Psychol.,* 1927, **39**, 367-389.
7. ————. The sense of feeling. In *The foundations of experimental psychology,* ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1929. Pp. 392-413.

[2]Since this study was completed a similar one has appeared dealing with warmth sensations (4). With increase in temperature, from 35 to 47 degrees, the number of points responding increased from 30 to 98 per cent in some cases (p. 59). The average increase was from 31 to 56 per cent.

8. TITCHENER, E. B. Experimental psychology. A manual of laboratory practice. Vol. I, Qualitative experiments. Part I, Students' manual. New York: Macmillan, 1901. Pp. xx+214.
9. ———. A textbook of psychology. New York: Macmillan, 1910. Pp. vii+565.
10. WHEELER, R. H. The science of psychology. New York: Crowell, 1928. Pp. xvii+556.

University of Nebraska
Lincoln, Nebraska

THE FUSION FREQUENCY IN DIFFERENT AREAS OF THE VISUAL FIELD: II. THE REGIONAL GRADIENT OF FUSION FREQUENCY*[1]

From the Department of Psychology of Yale University

ROBERT T. ROSS

The preceding paper in this series gave evidence of consistency of foveal fusion frequency in a given subject at any one experimental session, but the results seemed to point to increased dispersion of foveal fusion frequencies in the same subject from session to session and still greater dispersion when the mean foveal fusion frequencies for different subjects are compared.

If we assume that the physiological state of a given subject at one experimental session is constant, then the variability obtained during one such session is attributable to variation in judgment. We are forced to explain variations over and above this minimal amount, in the same or different subjects, as due to physiological differences. If we could by some method partial out the effects of these physiological differences we would expect the data obtained from different subjects to show much less variation. The present study attempts to do this by measuring the regional gradients of fusion frequency in different subjects.

If the physiological processes which change foveal fusion frequency affect the retina as a whole, then the differences in foveal fusion frequency should be a measure of the differences in physiological state over the entire retina. If, then, the fusion frequency at any point on the retina be expressed in terms of the foveal fusion frequency and the foveal fusion frequencies for different sets of readings are then equilibrated, the data obtained during different sessions should show more agreement after the transformation, since the effects of physiological differences have been partially eliminated. Fortunately,

*Accepted for publication by Raymond Dodge of the Editorial Board and received in the Editorial Office, June 24, 1935.

[1]This study is from a doctoral dissertation presented to the Faculty of Yale University in May, 1934.

this transformation is not difficult to accomplish, for if each foveal fusion frequency be taken as unity for a particular session, and the peripheral fusion frequencies obtained during this session expressed as percentages of this foveal value, we have accomplished the necessary conversion.

To date, the findings with regard to regional fusion frequency have been quite inconsistent. Most of the early workers, Aubert (2), Exner (9), Bellarminow (4), Sherrington (15), Polimanti (14), Allen (1), and Woog (16), have reported that the peripheral fusion frequency is higher than the foveal. On the other hand, Aubert (2), Exner (10, 11), Bellarminow (4), Brewster (7), and Braunstein (6) have reported the opposite effect. The more recent works of Lythgoe and Tansley (13), Creed and Ruch (8), and Hecht and Verrijp (12) have indicated that the relative magnitudes of foveal and peripheral fusion frequencies depend on several factors: (1) the brightness of the test-patch, (2) the adaptation of the eye, (3) the surrounds of the test-patch, and (4) the size of the test-patch. The question still remains, however, as to how the gradients of a large group of subjects would compare if these four variables were held constant. Using the normative group and the apparatus described in the preceding paper of this series the writer has investigated this problem.

Method

With a 40′ test-patch of 2.44 ml. viewed against dark surrounds after two minutes' light adaptation to an illumination of 3.42 ml. and one minute in the dark, the fusion frequency was determined at the fovea, 5°, 10°, 20°, 30°, and 45° along the nasal and temporal coordinates of the retina, and at the fovea, 5°, 10°, and 20° along the superior and inferior coordinates. The determination of one entire coordinate was made at one session according to the following scheme: (1) ten readings (five for each threshold) at the fovea, (2) ten readings at 5°, (3) ten readings at 10°, . . . and so on to (6) ten readings at 45°, (7) ten readings at the fovea (0°), (8) rest period, (9) ten readings at the fovea, (10) ten readings at 45°, (11) ten readings at 30°, . . . and so on to (15) ten readings at the fovea. This procedure balanced out fatigue effects (which, it might be said, failed to appear, the first ten reading showing remark-

able agreement with the later readings) and gave 40 readings of the foveal fusion frequency. Between each set of 10 readings the subject light-adapted for two minutes and then sat in the dark for one minute as previously described.

The Regional Gradient of Fusion Frequency

As would be expected from the previous study, the absolute values of the foveal fusion frequency differed from subject to subject so that when the absolute values of the fusion frequencies along the temporal coordinate are plotted against peripheral angle a situation shown graphically in Figure 1 is obtained. Even in Figure 1 some

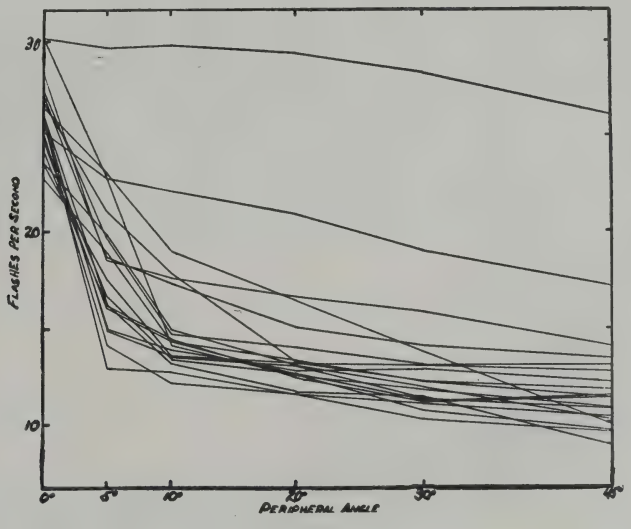


FIGURE 1

The Relation of Fusion Frequency to Peripheral Angle Along the Temporal Coordinate for 19 Normal Cases

general agreement is evident. To eliminate as far as possible the influence of whatever physiological differences contribute to individual differences in foveal fusion frequency, the gradients were expressed in percentage of foveal fusion frequency for each subject as shown in Table 1. When these results are plotted, the curves

TABLE 1

REGIONAL GRADIENTS OF FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE

Case	F.F.F. in flashes per sec.	5°	10°	20°	30°	45°
1*	30.05	68.0	57.2	50.5	47.1	——
2	28.32	66.0	62.3	59.0	55.8	49.3
3	24.67	65.2	57.4	54.1	49.4	45.8
4*	30.15	98.5	98.7	97.5	94.4	86.5
5*	26.47	87.5	72.0	62.6	52.4	37.4
6	24.97	56.4	48.6	46.4	40.8	38.0
7*	25.25	90.7	88.1	82.9	75.4	68.4
8	30.07	76.4	47.4	41.7	38.8	35.9
10	25.78	50.1	49.3	45.0	42.7	40.1
11	25.93	58.1	52.6	50.9	50.1	48.8
12	24.25	66.3	59.7	50.9	45.7	46.7
13	25.70	57.0	54.4	49.6	50.2	47.6
14	27.37	72.5	52.1	51.4	47.5	47.7
15	26.78	78.4	66.9	46.3	44.9	44.1
16	23.50	75.1	58.0	55.7	47.5	45.8
17	23.55	84.0	63.3	55.5	50.5	43.4
18	27.12	69.1	64.1	55.9	51.8	49.3
19	22.78	82.9	62.9	55.2	46.7	42.2
20	25.58	63.3	51.3	46.2	44.7	34.6
21	24.70	67.8	54.8	51.1	45.2	46.3
Mean	25.88	68.04	57.25	50.91	47.02	44.10

*Omitted from mean.

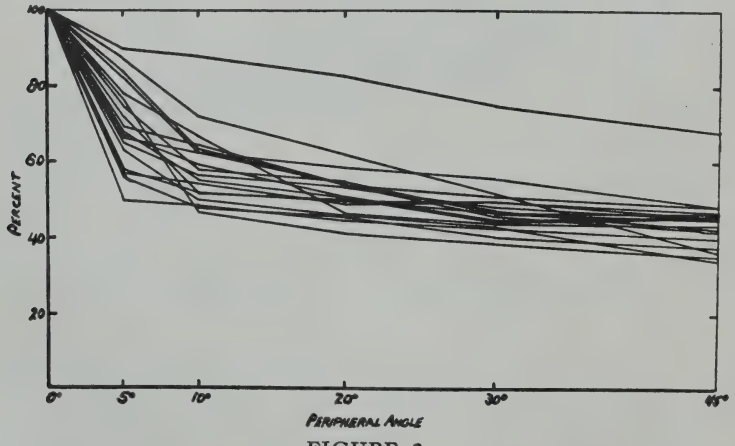


FIGURE 2

THE RELATION OF FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE TO PERIPHERAL ANGLE ALONG THE TEMPORAL COORDINATE FOR 18 NORMAL CASES

shown in Figure 2 are obtained. The mean gradient for the temporal coordinate is given for 17 normal cases in Figure 3.

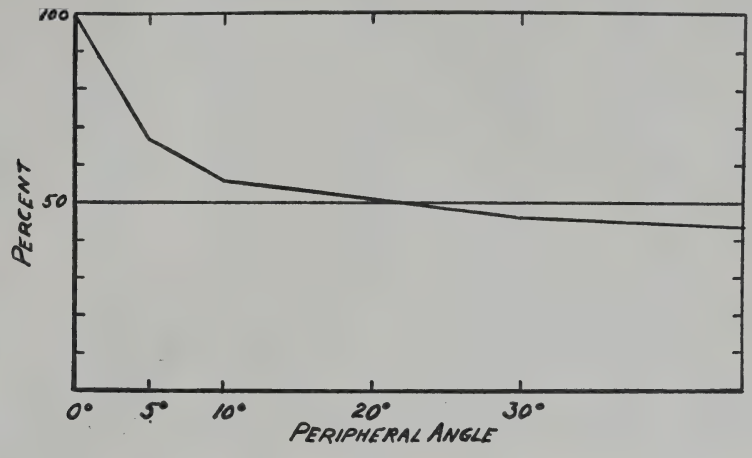


FIGURE 3

THE RELATION OF MEAN FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE IN PERIPHERAL ANGLE ALONG THE TEMPORAL COORDINATE FOR 16 NORMAL CASES

COMPARISON OF THE MAJOR COORDINATES

The data given above are based on temporal coordinate measures only. It is, of course, of value to have some comparison of the temporal with the nasal, superior, and inferior coordinates. Bellarminow (4) reported that the fusion frequency is higher in the nasal than in the temporal region. Braunstein (6) verified this result for the dark-adapted eye, but found a reverse relation for light adaptation. Hecht (12) found all four gradients practically alike.

In the present study the gradients for all four coordinates were measured for eight subjects. A graphical presentation of the results is given in Figures 4, 5, and 6. The results seem to bear out the findings of Bellarminow for this experimental situation, although the differences are hardly statistically significant. In general, however, we may accept the situation along the temporal coordinate as representative of the retina as a whole.

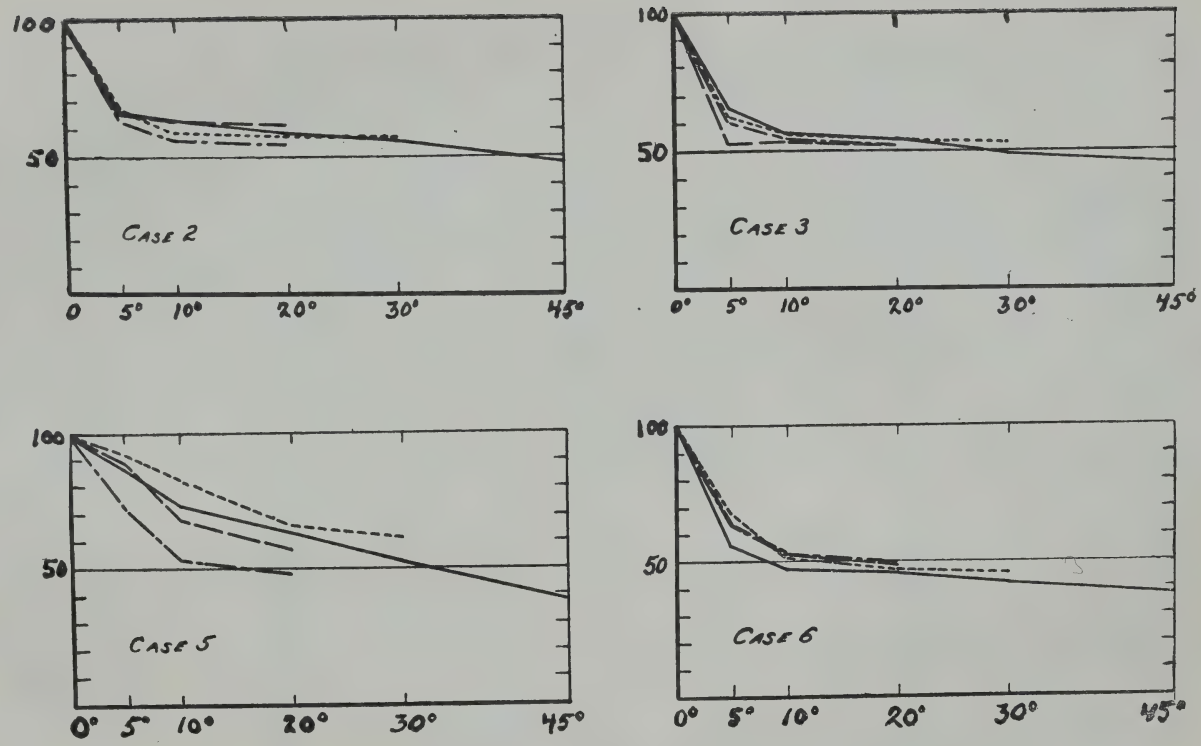


FIGURE 4

THE RELATION OF FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE TO PERIPHERAL ANGLE ALONG THE FOUR MAJOR COORDINATES OF THE RETINA

Solid lines refer to temporal coordinate, dotted lines to nasal coordinate, dashed lines to superior coordinate, and center lines to inferior coordinate.

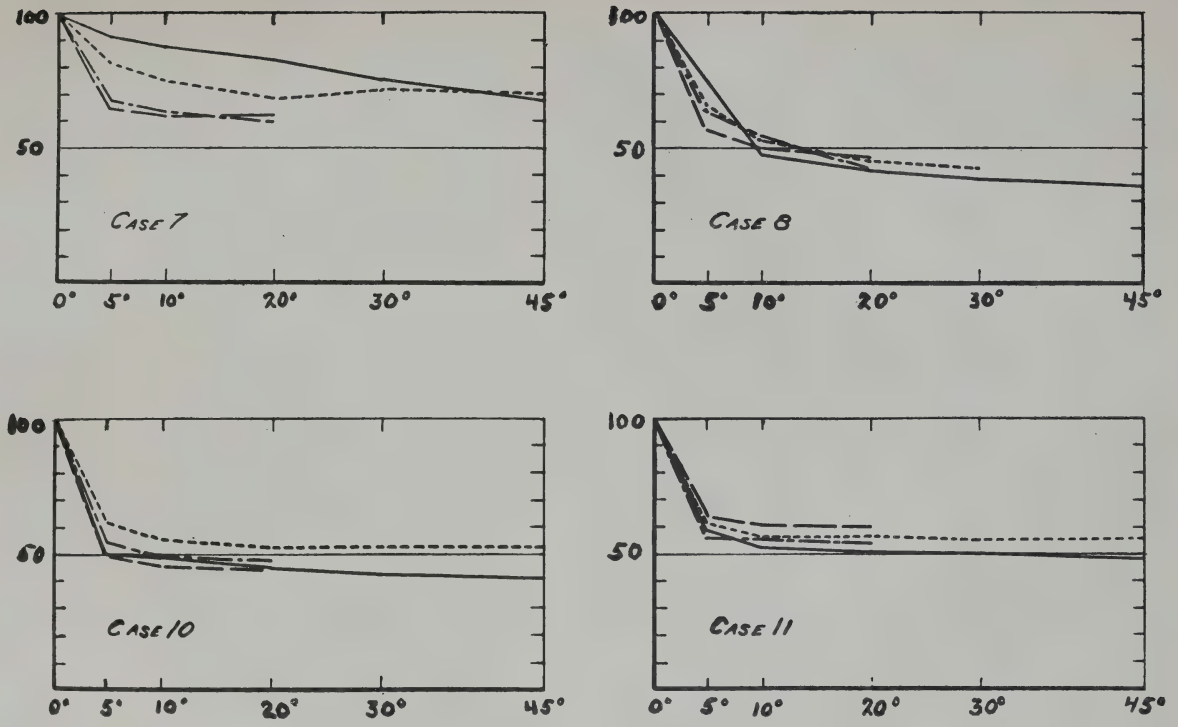


FIGURE 5

THE RELATION OF FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE TO PERIPHERAL ANGLE ALONG THE FOUR MAJOR COORDINATES OF THE RETINA
Solid lines refer to temporal coordinate, dotted lines to nasal coordinate, dashed lines to superior coordinate, and center lines to inferior coordinate.

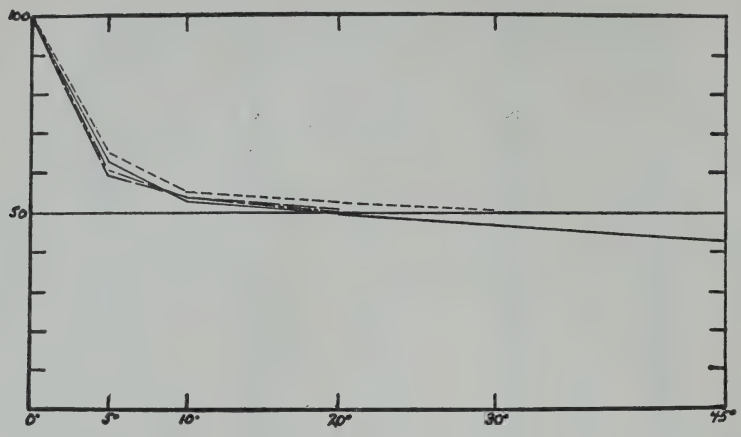


FIGURE 6

THE RELATION OF MEAN FUSION FREQUENCY IN PERCENTAGE OF FOVEAL VALUE TO PERIPHERAL ANGLE ALONG THE FOUR MAJOR COORDINATES OF THE RETINA, FOR SIX NORMAL CASES

Solid lines refer to temporal coordinate, dotted lines to nasal coordinate, dashed lines to superior coordinate, and center lines to inferior coordinate.

DISCUSSION

A comparison of Figures 1 and 2 bears out our assumptions regarding the physiological processes. When the effects of the physiological situation are partialled out, as far as possible, the gradients present a far less variable picture than they do in Figure 1. Figure 3 gives what we may consider to be a general picture of the regional gradient of fusion frequency in the average normal subject. To obtain this curve it has been necessary to discard three groups of data. Dangerous as this procedure is it may be shown that these cases evidence anomalous behavior. For the present it is sufficient to discard them on the basis of (1) their large variability of foveal fusion frequency values (see section I), (2) the lack of agreement in the various retinal coordinates, and (3) the criterion of "huge error."

We are thus left with a group of gradients which, if the physiological factor has been completely removed, should fall on one and the same line; or at least within the limits of individual variability. Obviously this is not the case. If we assume that we have partialled

out the adventitious changes in physiological condition which contribute to variability, then these residual differences must be due to fairly constant physiological conditions unique for each individual. In other words, it must be that the physiological process in any individual, when uninfluenced by transient conditions, follows a general pattern which is the same for all individuals but which differs in detail from one person to another.

This type of situation is well demonstrated in the work of Hecht and Verrijp (12) in which the writers find it possible to fit a general curve to their experimental data, but the constants of the equation which fits S. H.'s data are different from the constants in the formulation which fits C. V.'s.

The situation thus points to the possibility of formulating theories which will account for the general form of any visual function, but the evaluation of the constants in any such formulation will be unique for each case.

This is an excellent example of the fundamental difference between the tasks of physical and biological science. It is also evidence of the extreme care which must be used if theories are not to be built around exceptional cases.

Summary and Conclusions

The fusion frequencies of 20 normal subjects were determined at 0, 5, 10, 20, 30, and 45 degrees in the retinal periphery with a 40′ test-patch, illuminated at 2.44 ml., having dark surrounds and viewed after two minutes' adaptation to an illumination of 3.42 ml. and one minute of dark adaptation. Each value was determined by at least 20 individual readings. Under these experimental conditions it was found that:

1. Fusion frequency decreases with increase in peripheral angle.
2. "Normal" subjects have similar gradients.
3. A comparison of the major coordinates of the retina shows that the temporal gradient is representative of other retinal coordinates.
4. There is inconclusive evidence that the nasal gradient is generally higher than the temporal.
5. In no case did the regional gradient show a positive slope.

The results indicate that, when individual differences which may

be attributed to differences in physiological state are partialled out, both individual variability from session to session and individual differences in report tend to disappear. In other words, we have good reason to believe that visual process is fundamentally the same in different individuals and that the apparent differences are due to physiological factors which influence visual sensitivity.

References

1. Allen, F. Some phenomena of persistance of vision. *Phys. Rev.,* 1909, **28**, 45-56.
2. Aubert, A. Physiologie der Netzhaut. Breslau: Morgenstern, 1865. Pp. 394.
3. Basler, A. Ueber die Verschmelzung von zwei Nacheinander erfolgenden Lichtreizen. *Pflüg. Arch. f. d. ges. Physiol.,* 1912, **143**, 245-251.
4. Bellarminow, L. Ueber intermittirende Netzhautreizung. *Graefes Arch. f. Ophth.,* 1889, **35**, 25-49.
5. Braunstein, E. P. Beitrag zur Lehre des intermittirende Lichtreizes der gesunden und kranken Retina. *Zsch. f. Psychol.,* 1903, **33**, 171-206, 241-288.
6. ————. Zur Lehre von den Kurzdaurenden Lichtreizen der Netzhaut. *Zsch. f. Sinnesphysiol.,* 1923, **55**, 185-229.
7. Brewster, D. On the duration of luminous impressions on certain points of the retina. *Brit. Asso. Rep.,* 1858, **28**, Pt. II., 6-10.
8. Creed, R. S., & Ruch, T. C. Regional variations in sensitivity to flicker. *J. Physiol.,* 1932, **74**, 407-423.
9. Exner, S. Bemerkungen ueber intermittirende Netzhautreizung. *Pflüg. Arch. f. d. ges. Physiol.,* 1870, **3**, 214-240.
10. ————. Zur Kentniss von der Regeneration in der Netzhaut. *Pflüg. Arch. f. d. ges. Physiol.,* 1878, **16**, 407-409.
11. ————. Weitere Untersuchen ueber die Regeneration in der Netzhaut und ueber Druckblindheit. *Pflüg. Arch. f. d. ges. Physiol.,* 1879, **20**, 614-626.
12. Hecht, S., & Verrijp, C. Intermittent stimulation by light. III. The relation between intensity and critical fusion frequency for different retinal locations. *J. Gen. Physiol.,* 1933, **17**, 251-268.
13. Lythgoe, R. J., & Tansley, K. The adaptation of the eye: its relation to critical frequency of flicker. *Med. Res. Council, Great Britain, Spec. Rep. Ser.,* No. 134, 1929. Pp. 72.
14. Polimanti, O. Ueber die sogenannte Flimmer-Photometrie. *Zsch. f. Psychol.,* 1899, **19**, 263-283.
15. Sherrington, C. S. On reciprocal action in the retina as studied by means of some rotating discs. *J. Physiol.,* 1897, **21**, 33-54.
16. Wood, P. De la persistance variable des impressions lumineuses sur les differentes regions de la retine. *C. r. Acad. sci. Paris,* 1919, **168**, 1222-1224.

Yale University
New Haven, Connecticut

DYNAMIC AND TEMPORAL CONTROL IN MUSIC*[1]

From the Department of Psychology, the State University of Iowa

O. Irving Jacobsen

Part I. Dynamic Control in Music

The control of dynamic values is one of the four basic marks of musicianship, the others being control of pitch, time, and timbre. Therefore a measure of capacity for such motor control and achievement in performance is much needed.

Wickham (14) standardized a measure of the natural capacity for the control of intensity with a non-musical instrument, based on the assumption that the acquired motor skills used in performance on musical instruments should be eliminated. The aim of the present investigation was to devise a means of measuring achievement in the control of intensity on a musical instrument, and to determine the relationship between the capacity measurement and the achievement measurement. For the sake of brevity, the former will be referred to as a measure of capacity for intensity control, and the latter, a measure of achievement in intensity control; the former was measured on a non-musical instrument, and the latter on a musical instrument or voice.

Experimental Procedure

The capacity measure. This measurement was conducted in the manner described by Wickham. The apparatus is a tone dynamometer which produces a tone of constant pitch that can be increased or decreased in intensity by moving either one of two contact riders, one for the observer and one for the experimenter. The experimenter presents one of the 35 intensity steps of the scale as a

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, July 8, 1935.

[1]Grateful acknowledgment is made to Dean C. E. Seashore and Doctor J. Tiffin for guidance, counsel, and inspiration, and to Professor F. E. Kendrie and the members of the Iowa State University Symphony Orchestra for cooperation in this study.

standard to the blindfolded observer, who attempts to match this intensity by moving the contact rider. The steps of the scale are graduated from weak to very strong intensities.[2] The final score is the average error, usually for 50 trials.

The achievement test. A power level indicator, Type 586, General Radio Company, Cambridge, Massachusetts, with a carbon button microphone, was used on a six-volt direct current. The readings on the meter dial are in terms of decibels.

For voice performance, the vowel "a" as in "ah," was used. The wind instrument and voice observers performed directly into the microphone, over which a wire screen was placed so as to keep the instrument bell or the observer at a constant distance from the microphone during the test. For piano performance, the microphone was placed in a constant position against the sounding board.

Musical tones, even at the best, manifest irregular fluctuations in intensity to a surprising extent, both in finer pulsations and in larger crescendo and diminuendo performances. Although the meter needle registers these faithfully, no attempt was made to record them; but the plan to accept the dominant intensity of the tone at about its mid-point was adopted. This corresponds roughly to what the ear hears in practice. The test, as adapted, was divided into four exercises.

1) *Even intensity.* The observer was asked to produce a tone termed the standard, and to reproduce it immediately with the same intensity. Five trials were made at each of three pitch levels, each at three degrees of intensity, namely, weak, medium loud, and

---

[2]The tone dynamometer is built on the principle of a series of induction coils in which the number of turns in the secondary circuit increases in a geometric ratio so as to make the successive increments in loudness conform approximately with Weber's law. Measurement of the output reveals that from steps 10 to 30 there is a gradual increase in the magnitude of the difference between two successive steps. The increment is .5 db. at step 10, and 4.5 db. at step 30, increasing uniformly between these points.

This increasing magnitude of the increment with loudness seems to be justified by the fact that the average of all the records here reported tends to show approximately the same amount of error throughout the entire range of intensities. For the present purpose, it is not advisable to use steps below 10 or above 30. When the generating tone comes from the 60-cycle A.C. current, as here used, there is a disturbing roughness in the steps from 31 to 35. (Calibration of the tone dynamometer is shown in Appendix B of the dissertation on file at the Library of the State University of Iowa.)

loud. A medium pitch was chosen for the specific instrument or voice, and the other two pitch levels were taken a perfect fifth below and a perfect fourth above this medium pitch. For piano performance the pitch levels were replaced by a single tone, a three-tone chord played by the right hand, and a six-tone chord, three tones played by each hand.

2) *Least producible increase in intensity.* The observer was required to perform the second tone with the least producible increase in intensity over the standard. Trials were made for three pitches each with the three standard intensity degrees as in Exercise 1.

3) *Successive "least producible increases" in intensity.* Starting from a tone of weak intensity, the observer was required to perform the least producible increase in intensity cumulatively in ten successive steps.

4) *Sustained even intensity.* The observer was asked to sustain a tone at even intensity for a period of five seconds, at each of the three pitch levels and the three intensities. For piano performance, this procedure was replaced by a trill at the three degrees of intensity.

Preliminary practice trials were permitted for all four exercises so as to familiarize the observer with the test procedure. If an observer changed pitch in any trial, that trial was discarded. The score for each exercise was the average deviation from the standard tone in terms of decibels.

Fifty-five unselected observers were tested. Those who performed on a musical instrument were tested on that instrument. All others were tested by voice. Eleven observers were high school pupils, 24 were undergraduate, and 20 were graduate university students. Thirty-five of the observers had received musical training.

Interpretation of Results. The results of these two tests are given in Table 1. The scores for the achievement test are in terms of decibels; for the capacity test, in terms of steps on the dynamometer. The observers are grouped as trained and untrained in music.

The capacity test. The coefficient of reliability for the capacity test is $+.79$, P.E., $\pm.01$. The average score for all observers is 1.4 steps, and the standard deviation, .47.

The performance on the capacity test is slightly better for the

TABLE 1

RESULTS FOR THE TWO INTENSITY CONTROL TESTS

Observer	Instrument	Achievement test: Ex. 1 Db.	Ex. 1 S.D.	Ex. 2 Db.	Ex. 2 S.D.	Ex. 3 Db.	Ex. 3 S.D.	Ex. 4 Db.	Ex. 4 S.D.	Comp. score	Capacity test: Score	Capacity test: S.D.
						(*a*) *Musical instruments*						
Musically trained												
2	Trumpet	.1	.28	1.2	.34	1.1	.30	.5	.93	.7	.8	.85
2 (Untr.)	Voice	1.1	.92	2.2	.99	2.3	1.02	3.1	1.38	2.2		
4	Piano	.1	.28	1.5	.96	2.1	.72	.1	.04	1.0	.8	.64
4 (Untr.)	Voice	1.1	.86	2.2	.98	4.0	2.33	3.6	1.41	2.7		
9	Piano	.3	.46	1.4	.82	2.3	.69	.7	.59	1.2	1.0	.96
32	Clarinet	.4	.67	1.7	.84	1.8	.87	.8	.78	1.2	1.5	1.34
17	Clarinet	.3	.44	1.5	.75	1.2	.40	2.0	1.24	1.3	1.2	.97
20	Piano	.3	.42	1.6	.91	2.6	1.41	2.0	1.52	1.6	1.3	1.08
20 (Untr.)	Voice	1.3	.98	2.2	.94	2.7	1.58	2.6	1.42	2.2		
39	Clarinet	.5	.62	1.9	.81	1.8	.88	2.0	1.32	1.6	1.5	1.18
33 (Untr.)	Piano	.3	.32	1.9	.72	2.9	1.02	2.3	1.38	1.8	1.5	.96
33	Voice	1.4	1.08	2.8	1.47	4.9	2.30	5.1	1.66	3.6		
14	Fr. horn	.5	.58	1.7	.69	4.3	2.00	2.4	1.30	2.2	1.2	1.03
8	Trombone	1.2	1.34	2.3	1.33	2.8	1.64	4.2	1.63	2.6	1.0	.92
Average:	*inst. tr.*	*.4*		*1.6*		*2.3*		*1.8*		*1.5*	*1.2*	
	vc. untr.	*1.2*		*2.4*		*3.5*		*3.8*		*2.7*		
						(*b*) *Voice*						
16		.8	.76	1.7	.59	1.4	.45	1.8	1.36	1.4	1.2	.89
42		.6	.65	1.6	.58	1.6	.54	2.1	1.11	1.5	1.6	1.28
51		.9	.89	1.7	.35	1.6	.80	2.7	.78	1.7	2.1	1.59
27		.6	.72	2.0	.80	1.6	.80	2.8	.94	1.8	1.4	.95
38		.7	.30	1.6	.32	1.5	.67	3.6	.96	1.9	1.5	1.06
23		1.2	1.27	1.9	.73	1.6	.80	3.0	1.00	1.9	1.3	1.06
3		.8	.71	1.9	.62	2.0	1.34	3.4	1.17	2.0	.8	.90
10		1.1	.99	2.4	1.52	1.7	.96	2.6	1.12	2.0	1.1	.94
24		1.1	.73	2.4	1.22	2.5	.85	2.0	1.24	2.0	1.3	1.32
36		.6	.76	2.4	1.52	1.7	.71	3.5	.99	2.1	1.5	1.40
53		1.0	.88	1.8	.69	1.3	.41	4.1	1.42	2.1	2.3	2.18
18		1.1	1.18	2.0	.79	2.1	1.38	3.1	1.92	2.1	1.2	1.12
50		1.1	.97	2.4	1.09	2.9	1.37	1.9	1.19	2.1	2.0	1.80
6		1.3	.97	2.1	.84	2.0	.91	3.3	1.64	2.2	.9	.81
35		1.5	1.20	2.1	.81	1.8	.91	3.2	1.63	3.3	1.5	1.30

TABLE 1 (*continued*)

Observer	Instrument	Achievement test Ex. 1 Db.	Ex. 1 S.D.	Ex. 2 Db.	Ex. 2 S.D.	Ex. 3 Db.	Ex. 3 S.D.	Ex. 4 Db.	Ex. 4 S.D.	Comp. score	Capacity test Score	Capacity test S.D.
5		1.6	.68	2.3	.84	2.3	.80	3.1	1.21	2.3	.9	.98
12		.9	1.08	2.3	1.63	3.1	1.58	3.2	1.37	2.4	1.1	.98
28		1.6	1.36	2.3	1.22	2.3	1.28	4.0	1.89	2.6	1.4	1.23
47		.9	.71	2.4	.98	4.0	1.29	3.6	2.02	2.7	1.8	1.38
11		1.1	.85	2.7	.95	2.5	.91	4.8	1.87	2.8	1.1	1.25
7		2.0	1.71	2.6	1.48	2.7	.78	3.9	2.33	2.8	1.0	1.12
25		1.4	1.18	2.2	.92	3.8	1.17	5.8	1.31	3.3	1.3	1.11
54		1.3	.92	2.8	1.45	2.5	.96	6.6	1.86	3.3	2.6	1.80
40		1.3	.86	2.3	1.09	4.2	1.11	6.6	1.64	3.6	1.6	1.52
44		2.4	1.69	3.2	1.71	2.6	1.08	6.6	1.49	3.7	1.6	1.44
Average		*1.1*		*2.2*		*2.2*		*3.8*		*2.3*	*1.4*	
Untrained (All tested by voice)												
22		1.3	1.12	2.6	1.32	1.9	.94	2.9	1.26	2.2	1.3	1.29
37		1.6	1.42	2.3	1.18	2.6	1.43	3.6	1.33	2.5	1.5	1.31
13		1.5	1.37	2.5	1.10	2.5	1.32	3.9	2.00	2.6	1.2	.99
1		1.6	1.45	3.6	1.82	2.3	1.42	3.4	1.58	2.7	.5	.87
46		1.8	.96	2.2	1.16	2.3	1.41	4.6	2.11	2.7	1.7	1.54
43		1.7	1.31	2.0	.87	2.8	1.21	5.0	1.94	2.9	1.6	1.33
52		1.3	1.17	2.4	1.32	1.9	.70	6.1	1.29	2.9	2.2	2.08
29		2.4	1.56	3.5	1.77	3.2	2.13	6.2	1.57	3.8	1.4	1.32
55		3.4	1.15	4.3	1.43	4.5	1.50	5.6	1.76	4.4	3.2	1.89
High school boys with unchanged voices												
45		.9	.90	2.1	1.05	1.9	.83	3.1	1.45	2.0	1.7	1.65
21		1.4	1.06	2.4	1.22	2.7	.96	3.2	1.81	2.4	1.3	.85
31		1.3	.96	2.6	1.29	2.9	1.04	3.3	.96	2.5	1.4	1.25
30		1.0	.68	2.1	1.13	2.7	1.10	5.8	1.97	2.9	1.4	1.11
26		1.3	1.24	2.4	1.02	4.2	1.64	4.1	1.98	3.0	1.3	1.39
34		1.5	1.39	2.2	.88	4.3	1.32	3.6	1.83	3.1	1.5	1.35
High school boys with changing voices												
41		2.1	1.71	2.9	1.58	2.3	1.55	3.6	.94	2.7	1.6	1.30
15		2.6	1.49	3.1	1.52	2.5	.97	3.9	1.59	3.0	1.2	.89
48		2.4	1.60	3.6	2.29	3.1	2.02	6.6	2.49	3.9	1.8	1.46
19		3.0	1.98	3.8	1.71	2.6	1.57	6.2	1.68	3.9	1.3	1.14
49		2.4	1.50	3.1	1.31	5.5	2.96	5.6	1.83	4.2	1.8	1.35
Average (Untr.)		*1.8*		*2.8*		*2.9*		*4.7*		*3.1*	*1.6*	
Total average		*1.3*	*1.07*	*2.3*	*1.17*	*2.5*	*1.24*	*3.7*	*1.64*	*2.4*	*1.4*	*1.23*
Standard deviation		*.71*		*.6*		*.92*		*1.68*		*.69*	*.47*	

trained than for the untrained observers, the ratio being (35 cases) 1.4 steps to (20 cases) 1.6 steps, respectively. This is a fair validation of this test as a measure of capacity when training is disregarded.

The observers who are instrumentally trained do a trifle better than those who are vocally trained, the ratio being (10 cases) 1.2 steps to (25 cases) 1.4 steps. This may be due in part to the fact that the capacity test is more analogous to instrumental performance than to vocal performance. There is probably also a limited selection since dynamic control is more emphasized in instrumental than in vocal training.

Observer number 1, who was untrained musically, obtained the best score in this test. He was a speech instructor and was trained to observe different shades of intensity in speech. The next best score was obtained by the musically trained observer number 2, who had been a telegrapher. He stated that the various intensities, as heard at the different intensity steps, were associated with various code signals in telegraphy. Although this test does not involve the acquired motor skills used in musical performance, in the two cases just mentioned, it seems that a type of acquired skill not included in musical training did influence the scores obtained.

The capacity test of dynamic control deserves a place in the diagnosis of musical talent, particularly before a musical education is begun. Although this test does not directly involve motor control as required in the performance of any of the musical instruments or voice, it includes factors of memory and discrimination of intensity, combined with precision in voluntary motor control, all of which are greatly essential to musicians.

The achievement test. The data in Table 1 show that the standard deviations are, on the whole, satisfactory for measurements of this kind. The coefficient of reliability for the test is +.88, P.E., ±.01.

While each of the four exercises tested different elements in motor control of intensity, the composite score is given as an index figure. The average score is 2.4 db., and the standard deviation, .69.

The instruments seem to be divided into four groups according to dynamic control, namely, stringed instruments, wind instruments, the voice, and percussion instruments including piano. In each of

these four groups of instruments, dynamic performance demands a different set of muscles and skills. Even within the groups of wind, stringed, and percussion instruments, different sets of muscles and skills are involved in dynamic performance.

The tendencies in difficulty of intensity control in musical performance, as indicated by this investigation, are given in outline form below.

Characteristics of dynamic control

Untrained observers (20), 3.1 db.;	Trained observers (35), 2.1 db.
Vocally trained (25) 2.3 db.;	Instrum. trained (10), 1.5 db.
Inst. tr. but voc. tested (4), 2.7 db.;	Inst. tested (same 4), 1.3 db.
Boys' changing voices (5), 3.6 db.;	Boys' unchang. voices (6), 2.6 db.
Weak intensities (55),[3] 2.55 db.;	Strong intensities (55), 1.62 db.
Low pitches (55), 2.57 db.;	High pitches (55), 1.71 db.

Trained observers do better than untrained observers in the ratio of (35 cases) 2.1 db. to (20 cases) 3.1 db. This difference should be analyzed: first, because the measure of training here used is not quantitative; secondly, because the difference varies for voice and instrumental training; and, thirdly, because there is undoubtedly a marked selection on the part of those who have been trained; that is, those who have the natural capacity are likely to seek training.

The instrumentally trained do better than the vocally trained observers in the ratio of (10 cases) 1.5 db. to (25 cases) 2.3 db. This difference also needs further analysis: first, because the instrumental training had been far more intensive than the vocal training; and, secondly, because it seems probable that, conditions being equal, an instrumental test would be more precise than a vocal test.

Instrumental training does not improve dynamic control by voice as shown by the ratio of (4 cases) 1.3 db. in the instrumental test to (4 cases) 2.7 db. for the same individuals in the voice test. This fact throws doubt upon the validity of the difference stated above as possibly due to training, because the untrained were all tested by voice. These figures show not only that the instrumental training does not transfer to voice, but possibly also that the voice test, as a rule, should yield a larger error than the instrumental test.

Boys with unchanged voices do better than boys with changing

[3]The figures in parentheses indicate the number of observers concerned; the average scores are shown in decibels.

voices in the ratio of (6 cases) 2.6 db. to (5 cases) 3.7 db. Boys with changing voices complained that they could hear the differences but experienced difficulty in control of the voice.

The precision in control is better for the strong intensities than for the weak intensities in the ratio of (55 cases) 1.62 db. to (55 cases) 2.55 db., including both trained and untrained observers. Control for medium intensity is approximately equal to the average for all intensities.

The control is better for high pitches than for low pitches in the ratio of (55 cases) 1.71 db. to (55 cases) 2.57 db. The control for a medium pitch is approximately a mean between that for the high and the low pitches.

The achievement test of dynamic control can be used to determine the stage of development of a player in dynamic control on his instrument. Since it is an achievement test, it cannot be taken by other than trained observers, except for voice.

The coefficient of correlation between the capacity and the achievement test was found to be $+.34$, P.E., $\pm.08$, an indication that relationship is present, although it is low.

A Standard Achievement Test of Dynamic Control in Music. A standard test for the control of intensity on musical instruments should adequately measure skill in such control. The procedure suggested below is based on the findings of the present study. Since it was found that the intensity score for a medium intensity was approximately equal to the average score for the three degrees of intensity, and that the score for a medium pitch was approximately equal to the average for all three pitch levels, a medium loud tone at a medium pitch is used throughout the standard test.

Exercises 1 and 2 of the preliminary test are retained but with slight changes. Exercise 3, or "successive 'least producible increases' in intensity," is discarded because there is no demand for such performance in a musical situation. Exercise 4, or "sustained even intensity," also is discarded because of its similarity to Exercise 1, and because it is very difficult to record a valid score for such an exercise. The following two exercises of the standard test involve requirements which occur frequently in dynamic performance.

Exercise 1. Even intensity. The observer is asked to perform

the first phrase of "America,"[4] and to maintain all tones at the same degree of intensity. The intensity of the first tone becomes the standard. The phrase is repeated ten times. The intensities for the second and fifth tones only are recorded in terms of variations from the first or standard tone.

Exercise 2. Least producible differences in intensity. In one series of trials the observer is instructed to produce a standard tone of medium pitch and intensity, and then to reproduce it, pitch constant, with the least producible increase in intensity. In the other series, he is required to proceed in the same manner, except that he produces the least perceptible decrease in intensity. The trials are grouped as follows: first, ten trials of increased intensity, then ten trials of decreased intensity; this order is repeated five times. If the observer fails to produce a recordable difference, he is told that the trial is rejected and he is given a new trial.

Part II. The Motor Control of Rhythm and Intensity by Members of a University Symphony Orchestra

The present investigation was undertaken to determine: (1) the motor control of rhythm and intensity for each member of a symphony orchestra, and (2) the relationship between these two control tests and the Seashore sensory tests of rhythm, intensity, pitch, and timbre.

Apparatus and Procedure. The 73 members of a university symphony orchestra were given the two motor control tests individually, and the Seashore sensory tests were given as group tests, after the control tests had been taken by all members.

Control of rhythm. The rhythm meter as described by R. H. Seashore (11) was used. The apparatus maintains a continuous pattern as marked off by clicks in a telephone receiver, and the observer is required to keep time with this pattern by tapping a telegraph key. The action of the observer is recorded by a pen-tracing on a revolving paper disc; deviations from the true pattern are recorded in hundredths of a second. The rhythmic pattern,

[4]This particular phrase was chosen because the pitch of the first tone is repeated twice, and it is necessary to keep pitch constant in measures of intensity in terms of decibels.

a dotted quarter, an eighth, and two quarter note values, is repeated at the rate of 50 times per minute. From the record obtained, a score is computed from the deviations which are in five units, in hundredths of a second, as follows: (1) 1.33; (2) 1.72; (3) 2.28; (4) 3.03; (5) and larger, 4.02.[5] After a reasonable practice period, three trials of ten patterns each were recorded. The average deviation for the entire 120 taps or trials is the score for the test.

Control of intensity. The power level indicator set-up described in the preceding section was used to measure dynamic control. The apparatus and procedure were varied slightly according to the requirements of the various musical instruments. For tympani recording,[6] the horn of the microphone was placed above the tympani membrane, and the test was conducted by having the player perform a roll, rather than a single tap. For stringed instrument recording, the regular microphone was replaced by a button microphone tied securely to the back of the instrument. No instructions were given for bowing the stringed instruments, but a record was kept of the manner of bowing by each player. The wind instrument players were instructed to maintain their instruments in a constant position in relation to the microphone.

The observers were instructed to begin with a standard tone of medium intensity and of a prescribed pitch. For the wind instrument players, the pitch of this standard tone was an open tone of medium pitch according to the compass of the instrument. For the stringed instrument players, the standard pitch was B-flat on the A string, for both violin and bass viol; E-flat on the D string, for viola; and A-flat on the G string, for cello.

The test consisted of two exercises with procedure as prescribed for the standard achievement test described in the preceding section.[7] The score for the test was the average variation from the standard intensity in terms of decibels.

---

[5]The five units are marked off on the paper disc for each note of the rhythmic pattern; their values were computed by Nielson (7).

[6]Exercise 1 for the tympani player was the performance of a standard intensity and matching this intensity twice for each of the ten trials.

[7]Control of pitch was recorded simultaneously with the intensity measures. The Model VI tonoscope (9) was used for recording the pitch control; the least variation that can be recorded is 1/20 of a tone. The pitch con-

The Seashore sensory tests. The sensory tests of rhythm, intensity, pitch, and timbre were given to the group with procedure and precautions as described in the manual of instructions (8). The timbre discrimination test, designed by Tiffin and Lewis (12), is a recent addition to the Seashore test battery.

Experimental Results. The individual scores and standard devi-

TABLE 2

RESULTS FOR THE CONTROL AND SENSORY TESTS

		Control tests				Sensory tests							
		Rhythm		Intens.		Rhythm		Intens.		Pitch		Timbre	
Obs.	Instrument	*R.M.	S.D.	Db.	S.D.	R.S.	Rk.	R.S.	Rk.	R.S.	Rk.	R.S.	Rk.
1	Viola	3.3	1.28	1.41	.73	88	40	88	40	89	95	84	70
2	Violin	2.8	1.61	1.23	.58	82	85	89	45	84	65		
3	Violin	2.7	1.03	1.36	.84	80	80	86	30	83	60	83	65
4	Violin	2.6	1.20	1.74	1.04	78	75	86	30	85	70	64	65
5	Violin	4.1	1.71	2.42	1.56	84	90	86	30	90	95		
6	Violin	2.6	1.42	1.33	1.04	80	80	86	30	85	70	80	45
7	Violin	2.7	1.36	1.12	.59	72	50	85	25	90	95	84	70
8	Violin	2.7	1.13	1.26	.59	80	80	70	0	74	25	61	0
9	Flute	2.5	1.28	1.64	1.22	88	95	79	10	82	55		
10	Bassoon	2.4	1.02	2.33	1.57	88	95	82	15	83	60	85	75
11	Oboe	2.4	1.18	1.70	1.23	86	95	84	20	91	95	73	15
12	Cello	2.4	1.22	1.18	.64	90	99	89	45	90	95	81	40
13	Fr. horn	2.2	.98	1.22	.58	85	95	87	35	85	70	82	60
14	Violin	2.5	1.23	1.34	.72	86	95	87	35	84	65	87	85
15	Violin	2.4	1.12	1.46	.94	74	60	84	20	88	90	87	85
16	Violin	2.2	1.18	1.44	.73	88	95	85	25	88	90	95	99
17	Violin	3.1	1.35	1.01	.40	84	90	90	50	89	95	85	75
18	B. viol	2.8	1.56	1.03	.46	82	85	90	50	90	95	87	85
19	Viola	2.5	1.17	1.21	.50	80	80	87	35	88	90	93	99
20	Violin	3.2	1.34	1.07	.42	84	90	88	40	80	45	88	90
21	Cello	1.7	.68	.96	.11	94	99	93	70	99	99	85	75
22	Clarin.	2.7	1.33	1.82	1.08	86	95	85	25	86	75	88	90
23	Flute	1.9	.85	1.12	.58	90	99	86	30	88	90	73	15
24	Viola	2.5	1.27	1.34	.71	88	95	73	0	85	70	82	60
25	Violin	3.8	1.42	1.09	.42	78	75	85	25	91	95	84	70
26	Violin	3.2	1.41	1.70	1.27	82	85	80	10	86	80	90	95
27	Violin	3.1	1.39	.96	.22	82	85	92	65	90	95	86	80
28	Violin	1.8	.77	1.14	.48	98	99	94	80	86	80	98	99
29	Violin	2.5	1.40	1.07	.47	82	85	88	40	81	50	87	85
30	Tromb.	3.6	1.27	.86	.11	90	99	94	80	92	99	86	80
31	Oboe	2.0	1.04	1.32	.68	94	99	75	5	92	99	90	95

trol of the observers was superior; therefore low average deviations were obtained. Had the Model II tonoscope (10) been used, the measurement would have been finer since the readings are in terms of vibrations rather than fractions of a tone.

TABLE 2 (*continued*)

Obs.	Instrument	Control tests: Rhythm *R.M.	S.D.	Intens. Db.	S.D.	Sensory tests: Rhythm R.S.	Rk.	Intens. R.S.	Rk.	Pitch R.S.	Rk.	Timbre R.S.	Rk.
32	Tympani	2.9	1.62	1.05	.50	96	99	82	15	90	95	91	95
33	Violin	4.1	1.22	1.07	.43								
34	Violin	2.6	1.63	1.25	.56	84	90	86	30	90	95	92	95
35	Trumpet	2.2	1.33	.91	.18	90	99	76	5	86	80	83	65
36	Trumpet	3.2	1.31	.98	.18	78	75	91	60	91	95	90	95
37	Fr. horn	3.8	1.32	1.56	.83	76	70	84	20	82	55	80	45
38	Violin	4.2	1.22	1.35	.60	70	40	83	20	86	75	89	95
39	Fr. horn	2.0	1.00	.94	.19	92	99	94	80	86	75	83	65
40	B. viol	3.1	1.81	.95	.30	82	85	93	75	84	65	75	20
41	Trumpet	3.3	1.32	.98	.23	84	90	92	65	81	50	80	45
42	Violin	2.8	1.13	1.07	.40	80	80	90	50	89	95	91	95
43	Trumpet	2.8	1.21	1.02	.29	81	85	90	50	89	95	87	85
44	Fr. horn	3.6	1.29	1.68	.72	80	80	80	10	84	65	79	40
45	Bassoon	3.5	1.45	1.04	.37	78	75	89	45	60	15	78	35
46	B. viol	2.9	1.34	1.26	.53	80	80	87	35	86	75	78	35
47	Clarin.	2.5	1.31	1.34	.50	80	80	87	35	88	90	88	90
48	Clarin.	3.3	1.48	1.06	.35	96	99	86	30	96	99	75	20
49	Flute	4.0	1.60	1.91	1.33	74	60	85	25	83	60		
50	Violin	2.5	1.27	1.12	.51	82	85	86	30	82	55	87	85
51	Violin	3.8	1.78	.99	.35	78	75	88	40	84	65	85	75
52	Viola	4.4	2.06	.94	.11								
53	Tromb.	4.0	1.13	1.12	.97	78	75	89	45	79	40	86	80
54	Cello	2.9	1.50	.88	.09	90	99	92	65	93	99	81	55
55	Violin	2.6	1.24	.96	.27								
56	Flute	2.1	1.04	2.14	1.35	90	99	84	20	83	60	81	55
57	Cello	4.0	1.32	.91	.28	84	90	89	45	84	65	88	90
58	Violin	3.3	1.40	1.11	.60	80	80	86	30	85	70	94	99
59	Tromb.	2.5	1.28	.87	.09	88	95	91	60	85	70	84	70
60	Clarin.	4.0	1.19	1.07	.41	80	80	93	75	90	95	91	95
61	Violin	2.2	1.03	.96	.07								
62	B. viol	3.2	1.30	.94	.15	78	75	91	60	91	95	87	85
63	Tuba	3.6	1.41	1.23	.68	72	50	85	25	91	95	79	40
64	Violin	3.0	1.36	1.33	.56	80	80	89	45	80	45	86	80
65	Tromb.	2.0	.91	.96	.24	92	99	95	85	87	85		
66	Violin	2.8	1.20	.99	.41	94	99	86	30	90	95	82	60
67	Trumpet	3.4	1.22	1.78	.53	80	80	82	15	84	65	76	25
68	Violin	2.1	1.21	.97	.14	90	99	92	65	94	99	88	90
69	B. viol	1.6	.86	1.31	.78	96	99	90	50	90	95		
70	Violin	2.1	1.15	.81	.08	92	99	96	95	96	99		
71	Viola	2.2	1.18	.84	.09	94	99	97	95	99	99		
72	Viola	3.8	1.29	.86	.06	76	70	89	45	88	90		
73	Cello	2.7	1.52	1.16	.44	80	80	81	20	86	75	91	95
Average		*2.9*	*1.27*	*1.34*	*.54*	*86*	*86*	*87*	*41*	*84*	*70*	*84*	*69*
Standard dev.		.67		.33		6.7		5.1		6.2		6.6	

*R.M., score in rhythm control (in 100ths of a second); Db., score in intensity control; S.D., standard deviation; R.S., raw score; Rk., rank.

ations for the two motor control tests, and the individual scores and ranks for the sensory tests are presented in Table 2, together with the averages and standard deviations for the entire group.

Rhythm control. The orchestra members tapped the rhythmic pattern with an average score[8] of 2.9, and a standard deviation of .67. The average standard deviation for the individual scores, 1.27, indicates that these scores were somewhat scattered.

The coefficient of reliability for the rhythm control test is +.91, P.E., ±.01, which is similar to that obtained for this test by R. H. Seashore (11), namely, +.89 for 135 observers.

The difficulty of tapping rhythm apparently is not determined by the values of the notes themselves, but rather by the relationship of the particular notes in the rhythmic pattern. The average errors for the four notes of the rhythmic pattern for the entire orchestral group are distributed as follows: dotted quarter note, 2.7; eighth note, 3.3; quarter note, 3.4; and quarter note, 1.8. The quarter note usually is considered the standard both in musical and experimental rhythmic performance. Notes longer or shorter than the standard note are followed by larger errors in the rhythmic tapping. According to the above results, the third note of the pattern, a quarter note, was the most difficult, and the fourth, also a quarter note, was the least difficult in rhythmic performance. An eighth, the second note, preceded the third or quarter note, and it was perhaps the short period of time between these two notes that caused greater difficulty in tapping the third note. A dotted quarter, the first note, preceded the eighth note, the second note of the pattern; the latter note was the second most difficult note in rhythmic performance, but in this case it was perhaps the long period of time between these two notes that caused greater difficulty in tapping the second note.

Although the observers varied in the type of errors made in

[8]Separate scores were computed for each of the four notes of the rhythmic pattern, and the final test score was the average of these four scores. Nielson (6) computed the measurement for only one of the four notes as the score for the test, but the inadequacy of such procedure is shown in the present investigation in which the average scores for the entire group varied considerably for the four different notes. Observer 19 had extremes as follows: dotted quarter note, 1.6; eighth note, 3.4; quarter note, 3.3; and quarter note, 1.8.

rhythmic action, there was a constant tendency for most of them to anticipate the rhythm, or to tap the four notes of the pattern early rather than late. This is shown by the fact that the ratio of anticipated to delayed taps was approximately 4 to 1, while the size of the errors was approximately the same for each. Remedial training in rhythmic performance for members of this group[9] should stress the avoidance of anticipated tapping.

Intensity control. The ability of the orchestra members to control intensity on their respective instruments is shown by the average score of 1.3 db., in Table 2. This is slightly better than the average score obtained in the preliminary test for the ten instrumentally trained observers, that is, 1.5 db. This slight difference may be due both to the fact that the orchestral group of observers had been trained more thoroughly than the other group, and to the small sample in the smaller group.

The coefficient of reliability for the intensity control test is +.85, P.E., ±.02, as compared to that of +.88 obtained in the preliminary test with 55 observers.

The standard deviation for the intensity control test is shown in Table 2 to be .33, an indication that scores of the orchestra members were quite concentrated.

For the matched intensity performance required in Exercise 1, there was a slight tendency toward increasing the intensity, the ratio for the number of trials increased and decreased being approximately 4 to 3. However, errors were greater when the intensity was decreased. Decreased intensity errors averaged .21 db.; increased intensity errors, .15 db.

Both the number and the size of errors were greater for the fifth than for the second tone of the phrase from "America," which constituted the even intensity requirement of Exercise 1. The average error and the number of errors were twice as great for the fifth as for the second tone of the phrase. The probable reason for this

[9]Ten of the observers had been or were conductors of band or orchestra, and each of these ten had only an average score in the rhythm control test. Perhaps the reason for this low ranking was that these observers were accustomed to varying tempos with retardations and accelerations as they desired, and to avoiding mechanical exactness in beating time. Although they understood the exacting requirements of the test, their experience may unconsciously have influenced their rhythmic action.

difference is that of the proximity of these two tones to the standard tone, and that two pitches differing from the standard tone preceded the fifth tone.

A decrease in intensity was found to be more difficult than an increase in intensity performance. The averages for Exercise 2 were 1.1 and 1.4 db., for the increased and decreased intensity, respectively.

Although no instructions for bowing the stringed instruments were given, the manner of bowing was recorded for each observer. If an observer asked about bowing, he was told to bow the exercises as he preferred. It was found that the bowing of the pairs of tones of Exercise 2 by the 46 observers, regardless of increased or decreased intensity requirements, was as follows: down and up bow, 33; up and down bow, 7; two half bows, down, 3; two half bows, up, 1; and up and down for increased, down and up for decreased intensity, 2. It is probable that an efficient performer has the skill to bow in any of the five manners mentioned and yet have excellent intensity control, but the fact that 16 observers of the string group, who received the best scores, used the down and up method of bowing for each pair of tones, indicates that this method is either superior to others, or it is the automatic method of bowing for observers who have the best intensity control. All of the remaining 30 stringed instrument players ranked in the lower 50 per cent in intensity control.

Marked timbre changes occurred with a change of bow stroke for most stringed instrument observers, according to the judgment of the experimenter who had had sufficient training to warrant such judgments. It was apparent that timbre changes interfered with, rather than aided intensity control. When two bow strokes were used in Exercise 2, the timbre change with the change of bowing was unusually prominent for more than 60 per cent of the observers. The stringed instrument observers who had the best intensity control, according to the test scores, had no noticeable change in timbre, regardless of bowing or intensity requirements of the test.

This timbre change with change of bow stroke was prominent when there was no change in intensity or pitch, and also when either or both were changed. The timbre and intensity relationship suggests the illusions which occur in musical performance, namely, that

the observer may perceive a change in a musical tone, but he cannot state whether it is a change in intensity, pitch, or timbre. The ensemble effect of a symphony orchestra must naturally be greatly altered if members of the stringed instrument groups, while using the same bow strokes, have these great individual changes of timbre. Perhaps it is the ability to maintain a similar or nearly constant timbre, regardless of bowing, pitch, or intensity, that distinguishes the great symphony orchestras from others.

Since pitch control was recorded, it was possible to determine the tendencies in pitch performance with certain intensity variations. When a constant intensity was required, there was a tendency to perform sharp in pitch more often than flat. Miles (6) also found that there was a tendency for his observers to sharp errors rather than to flat them.

If pitch errors occurred at all when observers increased the intensity in Exercise 2, 96 per cent of such errors were sharp, whereas if pitch errors occurred when observers decreased the intensity in Exercise 2, 98 per cent of such errors were flat. The performance on the clarinet was the only exception to this tendency toward sharping and flatting of pitch with increased and decreased intensity, respectively. When the requirement was a decreased intensity, clarinet players tended either to make no error or to sharp the pitch, whereas, when it was an increased intensity, the error in pitch was flat. This characteristic has been observed by Langenus (4) and others who have investigated the wood-wind instruments.

The Seashore sensory tests. The individual scores and ranks for the four sensory tests, given in Table 2, together with the number of orchestra players who took each test, show that all of these measures were high, due no doubt to the fact that this is a selected group of musicians. The homogeneity of the group is further shown by the low standard deviations for each of the four tests.

The relation between the various tests. The intercorrelations of the two motor control and the four sensory tests are given in Table 3. With the exception of the intensity discrimination test, the intercorrelations with intensity control and other tests are relatively low. Likewise, all intercorrelations for rhythm control, excluding rhythm discrimination, are relatively low. Timbre discrimination, which is a new test, correlated highest with intensity

TABLE 3
INTERCORRELATIONS OF THE CONTROL AND SENSORY TESTS

	Pitch discrim.	Intens. discrim.	Rhythm discrim.	Timbre discrim.	Rhythm control
Intensity control	.19	.52	.11	.11	.09
Rhythm control	.18	.05	.65	—.01	
Timbre discrim.	.29	.33	.16		
Rhythm discrim.	.28	.24			
Intens. discrim.	.34				

discrimination. There is a relationship between these two tests, since in timbre discrimination the observer must discriminate intensity differences of overtones which influence timbre changes.

An interesting by-product of this study is the observation that with the highly selected group of musicians here used, the intercorrelations between pitch, intensity, and rhythm discrimination are similar to those found in other studies made with unselected adult observers. Intercorrelations have been computed for 210 college students by Weaver (13); for 100 adults by Larson (5); and for 108 high school juniors and seniors by Brown (1). The correlation between pitch and intensity discrimination was found to be +.32 by Weaver, +.33 by Larson, +.25 by Brown, and +.34 in the present study. For intensity and rhythm discrimination, the correlation was found to be +.21 by Larson, +.23 by Brown, and +.24 in the present study. For pitch and rhythm discrimination, the correlation was found to be +.30 by Larson, +.11 by Brown, and +.28 in the present study. These results substantiate the belief that the Seashore tests measure rather stable and quite discrete musical powers.

The coefficient of correlation for the motor controls and their corresponding sensory tests, for the members of the three sections and for the entire orchestra group, are given in Table 4. The

TABLE 4
CORRELATION OF THE CONTROL AND SENSORY TEST RESULTS

	Intensity control and intensity discrim.	Rhythm control and rhythm discrim.
String section	.47	.69
Wood-wind section	.35	.48
Brass section	.54	.60
Entire orchestra	.52	.65

product-moment method was used for computing the coefficients for the scores for the entire orchestra group and for the string section. Because of the small number of cases in the brass and wood-wind sections, the Spearman Foot-Rule method of computing the coefficient was used for them.

All but one of these values is within the range called "substantial or marked relationship" by Garrett (2). R. H. Seashore (11) found a correlation of +.64 between rhythm control and rhythm discrimination, with 135 observers. Henderson (3), however, obtained a correlation of only +.37 between these tests, with 47 observers.

The conclusion to be reached from a study of the table is that sensory capacity and motor capacity, in the same musical factor, are not intimately related. The fact that an individual can discriminate very small differences in intensity is no indication that he has the motor skill to reproduce these same differences. A violin and a bass viol player took the intensity control test on the trumpet also. The scores were very much better for the stringed instrument than for

TABLE 5

CONVERSION OF RAW SCORES OF THE RHYTHM AND INTENSITY CONTROL TESTS INTO CENTILE RANK

Rhythm		Intensity	
Centile	Score in A.D.	Centile	Score in Db.
99	1.58	99	.81
95	1.91	95	.86
90	2.04	90	.91
85	2.15	85	.95
80	2.39	80	.96
75	2.43	75	.98
70	2.52	70	1.02
65	2.56	65	1.05
60	2.63	60	1.07
55	2.73	55	1.11
50	2.80	50	1.12
45	2.90	45	1.16
40	3.05	40	1.23
35	3.12	35	1.26
30	3.24	30	1.33
25	3.31	25	1.34
20	3.45	20	1.36
15	3.61	15	1.56
10	3.84	10	1.74
5	4.02	5	2.14

the trumpet test in each case. Since motor control varies greatly with the instrument on which the test is taken, it is impossible to predict outright, the ability of control, without a consideration of other factors such as interest, general motility and dexterity, physical development, age, training, and the like.

Centile rank for the two control tests. From the raw scores of the rhythm and the intensity control tests, tentative norms for musically trained observers were constructed; these are presented in Table 5.

Test results for members according to instruments and sections. The averages for motor control of rhythm and intensity, and for the four sensory tests, for players of each group of instruments, and of each section of the orchestra, are given in Table 6.

There is no great difference in averages in rhythm control of the players according to instruments or sections. This seems to indicate that training does not greatly influence rhythmic action, because

TABLE 6

SCORES FOR THE MOTOR CONTROL AND SENSORY TESTS FOR MEMBERS ACCORDING TO INSTRUMENTS AND SECTIONS

Instrument	No. of instru.	Control tests		Sensory tests			
		Rhythm	Intens.	Rhythm	Intens.	Pitch	Timbre
String section							
1st violin	14	3.0	1.4	83	84	88	87
2nd violin	16	2.8	1.2	83	87	85	84
Viola	6	3.1	1.3	84	87	90	80
Cello	5	2.8	1.2	86	89	89	85
Bass viol	5	2.8	1.2	84	90	88	82
Entire section	46	2.9	1.3	84	87	87	85
Brass section							
Trumpet	5	3.0	1.2	83	87	86	83
French horn	4	2.9	1.5	83	86	84	83
Trombone	4	3.0	1.1	87	92	86	85
Tuba	1	3.6	1.4	72	85	91	79
Entire section	14	3.0	1.3	83	88	86	83
Wood-wind section							
Flute	4	2.6	1.8	85	83	84	77
Oboe	2	2.2	1.6	90	80	91	82
Clarinet	4	3.1	1.4	86	88	90	86
Bassoon	2	2.9	1.8	83	86	72	82
Entire section	12	2.8	1.7	86	86	85	82
Percussion section							
Tympani	1	2.8	1.2	86	92	85	82
Entire orchestra	73	2.9	1.3	86	87	84	84

rhythmic performance on the various orchestral instruments varies greatly according to the performance possibilities on the instrument. Likewise, there is little difference in averages in intensity control of players according to instruments or sections, although the wood-wind section was slightly lower than the others.

The averages for the four sensory tests, for members according to instruments and sections, are similar to those for the entire orchestral group, and again indicate that the measures in musical talent are high for this group of selected musicians, and also that the four sections share equally in this talent.

References

1. Brown, A. W. The reliability and validity of the Seashore tests of musical talent. *J. Appl. Psychol.*, 1928, **12**, 468-476.
2. Garrett, H. E. Statistics in psychology and education. New York: Longmans, Green, 1926. Pp. 149-220. (See esp. p. 320.)
3. Henderson, M. T. Remedial measures in motor rhythm as applied to piano performance. Ph.D. Thesis, Univ. Iowa, 1931.
4. Langenus, G. Six articles for wood-wind players. New York: Fischer, 1923. (See esp. pp. 8, 19.)
5. Larson, R. C. Studies on Seashore's "measures of musical talent." *Univ. Iowa Stud.*, 1930, **2**, No. 174. Pp. 28.
6. Niles, W. R. Accuracy of the voice in simple pitch singing. *Psychol. Rev., Monog. Suppl.* (Iowa Stud.), 1914, **16**, 13-66.
7. Nielson, J. T. A study of the Seashore motor-rhythm test. *Psychol. Rev., Monog. Suppl.* (Iowa Stud.), 1930, **40**, 74-84.
8. Seashore, C. E. The tonoscope. *Psychol. Monog.*, 1914, **16**, No. 69, 1-12.
9. ————. Manual of instructions and interpretations for measures of musical talent. Chicago: Stoelting, 1919. Pp. 16.
10. ————. The measurement of pitch intonation with the tonoscope in singing and playing. *Univ. Iowa Stud.*, 1930, **1**, No. 172, 7-29.
11. Seashore, R. H. Studies in motor rhythm. *Psychol. Rev., Monog. Suppl.* (Iowa Stud.), 1926, **36**, 142-189.
12. Tiffin, J., & Lewis, D. Unpublished research conducted in the Iowa Psychological Laboratories.
13. Weaver, A. T. Experimental studies in vocal expression. *J. Appl. Psychol.*, 1924, **8**, 23-51.
14. Wickham, D. E. Voluntary control of the intensity of sound. *Psychol. Rev., Monog. Suppl.* (*Iowa Stud.*), 1922, **31**, 260-267.

286 Elmhurst Avenue
Elmhurst, Illinois

SHORT ARTICLES AND NOTES

RELATION OF SUGGESTIBILITY TO OTHER PERSONALITY TRAITS

An Attempt to Measure the Relation between Suggestibility and Unhealthy Mental Traits in Psychopathic and Normal Individuals

MARION R. BARTLETT

INTRODUCTION

In this study we have interested ourselves in the problem of the relation of direct waking suggestion to other measurable personality traits. A very considerable amount of controversy has raged around the notion of "suggestibility" as indicative of a pathological condition. For many decades it has been the theory of the Charcot school that suggestibility is closely linked with neurotic tendency, that it indicates a pathological condition (1). But little work other than that of clinical observation has been done to investigate the relation of suggestibility to mental disease. However, one such investigation has been reported, that of Davis and Husband (2). They used a scale of hypnotic susceptibility as a measure of suggestibility and the Thurstone Personality Schedule as a measure of maladjustment or neurotic tendency. Their correlation coefficient with the total score was .045, indicating no relation between suggestibility and neurotic tendency. In a previous article of our own (3) in which we used diagnosis of psychoneurosis in a mental hospital as the basis for neurotic tendency, we found that psychoneurotic subjects showed themselves to be no more suggestible than were the normal control subjects, thus corroborating the work of Davis and Husband.

In the present study we proposed further to investigate the problem by ascertaining the relation between direct waking suggestibility and a series of personality tests which purport to measure mental health. As a measure of suggestibility we have used the Hull technique of postural sway (4, pp. 50-51); and as a measure of other personality traits we have given the Bernreuter Personality Inventory and the Maller Character Sketches. While such techniques of measuring personality traits are far from perfect instruments, their use in attacking such a problem seems justified in their superiority over uncontrolled clinical observations from scattered sources.

The measurement of suggestibility, which in the early studies seemed to be in a hopelessly chaotic and controversial state, has been clarified and organized in the postulates set forth by Hull (4). He differentiated between the studies of *indirect* non-prestige suggestion and direct walking sug-

gestion, which he considered synonomous with susceptibility to hypnosis. He states (4, p. 385): "The writer is proceeding on the tentative assumption that . . . non-prestige suggestion is essentially different from—indeed quite without significant relation to—hypnosis and direct prestige suggestion in general." And elsewhere (4, p. 101): "It has been found wherever the question has been investigated that all forms of direct verbal or prestige suggestion correlate positively and to a rather high degree with each other and with susceptibility to hypnosis, whereas there is some indication that they correlate to a low or zero degree with indirect suggestion." This formulation throws considerable light on the previous experimental work with suggestion, the results of which did not seem to be in accord with one another, and which showed considerable ambiguity (5).

Present Study

1. *Procedure.* In the present study we used a group of 16 psychoneurotics, 16 dementia-praecox cases, and 20 normals as our subjects, making a total of 52 subjects. We made separate correlations with the smaller groups as well as with the total number, although the numbers were too small to give very significant results. This we did in order to observe whether or not there might be *tendencies* in the individual groups which were not observable when all were pooled together.

The subjects were taken to the laboratory and blindfolded. As this was being done a pointer was slipped under the blindfold. The subject was led to a certain spot at right angles to the wall and requested to stand in an easy relaxed position. The experimenter allowed one minute to elapse before starting the test, then proceeded to give it in standardized form (4, p. 50-51):

> Now you will begin to feel yourself falling forward. You are beginning to fall forward. You are falling forward. You feel yourself falling forward a little more; a little more; you are falling forward. You are falling forward over onto your toes. . . .

The experimenter continued the suggestion for a period of three minutes using a stop watch to measure the time. Records were taken by observing the point at which occurred the maximum amount of sway, as indicated by the shadow cast by the pointer on a chart on the wall. This chart was so constructed that it allowed for forward or backward sway from a central zero point scaled in degrees.

On a different day the subject was requested to fill out the personality inventories above described. The Bernreuter test was scored for neurotic tendency, self-sufficiency, introversion, and dominance; and the Maller test was scored for habit pattern, self-control, social adjustment, personal adjustment, mental health, readiness to confide, and the total score.

Results

Examination of the results in Tables 1, 2, and 3 indicate for the most part almost no relationship between suggestibility and the personality traits measured. Traits such as "neurotic tendency," "self-sufficiency," "self-control," "social adjustment," "personal adjustment," etc., which might very conceivably be closely allied to a tendency to react to environmental

TABLE 1

Correlation of Hull's Test of Suggestibility with Personality Tests

Test	All subjects grouped r	P.E.
Bernreuter		
*N*1 Neurotic Tendency	.050	.10
*S*2 Self-sufficiency	.006	.10
*I*3 Introversion	—.171	.09
*D*4 Dominance	—.064	.10
Maller		
Habit Pattern	.32	.09
Self-Control	.11	.09
Social Adjustment	.11	.09
Personal Adjustment	.01	.10
Mental Health	—.02	.10
Readiness to Confide	.27	.09
Total Maller	.09	.09

TABLE 2

Correlation of Hull's Test of Suggestibility with Personality Tests (When the subjects are divided into groups according to diagnostic category.)

Test	Psychoneurotic r	Dementia praecox r	Psychopathic r	Normal r
Bernreuter				
*N*1 Neurotic Tendency	.14	.12	.17	—.09
*S*2 Self-Sufficiency	—.18	.09	—.10	.04
*I*3 Introversion	—.013	—.17	—.02	—.29
*D*4 Dominance	—.36	.11	—.15	—.14
Maller				
Habit Pattern	.38	.38	.40	.017
Self-Control	—.13	.24	.03	.17
Social Adjustment	.34	—.14	—.15	—.11
Personal Adjustment	—.23	.09	—.10	.08
Mental Health	—.33	.08	—.18	.16
Readiness to Confide	.34	.28	.30	.19
Total Maller	—.10	.018	.07	.09

TABLE 3

SIGNIFICANT OR NEARLY SIGNIFICANT CORRELATIONS FROM TABLES 1 AND 2

Group	Test	r	P.E.$_r$	Chances in 100
All subjects	Habit Pattern	.32	.09	99
" "	Readiness to Confide	.27	.09	98
Psychoneurotic	Dominance	—.36	.15	96
"	Habit Pattern	.38	.15	97
"	Social Adjustment	.34	.15	95
"	Mental Health	—.33	.15	93
"	Readiness to Confide	.34	.15	94
Dementia praecox	Habit Pattern	.38	.15	97
" "	Readiness to Confide	.28	.16	88
Psychopathic	Habit Pattern	.39	.10	100
"	Readiness to Confide	.30	.11	97
Normal	Introversion	—.29	.14	91

stimuli without critical evaluation, failed to show any relation to our test of suggestibility. This result is in close agreement with the study of Davis and Husband, and our own previous study. It would appear that the theory that direct waking suggestibility is related to a pathological condition is not substantiated by the experimental evidence. It would rather appear to have very little relationship whatever to these other personality traits. What little relationship there is would seem from our study to be of a positive rather than a negative nature, as will subsequently be shown.

Throughout the series of correlations only two of the tests consistently showed coefficients sufficiently high to be significant. These were the correlations of suggestibility with Maller's "habit pattern," and "readiness to confide." "Habit pattern" essentially measures well-adjusted social habits. Below is a sample of the type of question which it includes.

1. This person is always wide awake; does not mislay or misplace things.
3. Insists on having his own way and likes to command and rule everybody.
8. Exaggerates and fusses about ordinary things and happenings.
10. Courteous; thoughtful of the welfare of others.

The questions were answered on the basis of being the "same" or "different" from an hypothetical person, and the scoring done on such a basis that a high score indicates a high degree of socially acceptable habits. In the case of this test then a limited amount of relationship seemed to exist between direct waking suggestion and habits of behavior which are socially adapted.

What "readiness to confide" measures is obvious in the title. A sample question is the following.

1. When in trouble the person does not like to suffer in silence.
   He feels much better after talking it over with somebody.

That such a correlation was found to exist between suggestibility and readiness to confide seems but little surprising in the light of the social requisites in each situation. A person socially suggestible would appear to be more ready to respond in a social situation by confiding to others his own problems and interests.

In addition to these two correlations there were a few of the correlations in the smaller groups which were nearly significant, although not high enough to have any predictive value. We shall indicate these by the separate groupings.

1. *Psychoneurotic Group.* Table 3 indicates that in the psychoneurotic group there is some tendency for a relation to be shown between the following: (1) suggestibility and *lack* of dominance; (2) suggestibility and habit pattern; (3) suggestibility and social adjustment; (4) suggestibility and *lack* of mental health; and (5) suggestibility and readiness to confide.

2. *Dementia-Praecox Group.* In this group only the two tests above indicated as consistently showing a relation to suggestibility were sufficiently high to be indicative—habit pattern and readiness to confide.

When the two psychopathic groups were combined only those two correlations common to both remained significant, those between suggestibility and habit pattern, and readiness to confide.

3. *Normal Subjects.* In the normal group there was a slight indication of a negative relation between suggestibility and introversion. None of the Maller tests were sufficiently high to indicate a relation. Six of the seven coefficients, however, were low positive values.

Summary

1. An investigation of the relation of suggestibility to a battery of other personality tests by the technique of correlation failed in most instances to reveal any relationship. From the results of our study it would appear that there is little relation between suggestibility and neurotic tendency, and other mentally unhealthy traits.

2. The two tests which did appear to have some positive relation to suggestibility were Maller's habit pattern and readiness to confide. These coefficients were statistically significant, though not sufficiently high to have predictive value.

3. In addition there appeared to be a *tendency* for a relation to ex-

ist in the psychoneurotic group between suggestibility and *lack* of dominance, *lack* of "mental health,"[1] and social adjustment; and in the normal group between suggestibility and extraversion.

References

1. Janet, P. Névroses et idées fixes. Vol. I. Paris: Ballière, 1904.
2. Davis, L. W., & Husband, R. W. A study in hypnotic susceptibility in relation to personality traits. *J. Abn. & Soc. Psychol.*, 1931, **26**, 175-182.
3. Bartlett, M. R. Suggestibility in psychopathic individuals. A study with psychoneurotic and dementia praecox subjects. *J. Gen. Psychol.* (In press.)
4. Hull, C. L. Hypnosis and suggestibility. New York: Appleton-Century, 1933. Pp. xii+416.
5. Murphy, G., & Murphy, L. B. Experimental social psychology. (See esp. Ch. 4: The learning process in social situations.) New York: Harper, 1931. Pp. 709.

Yale University
New Haven, Connecticut

TYPES AND METHODS OF LEARNING

Carl N. Rexroad

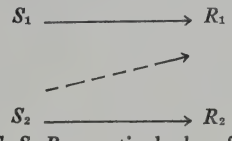
The techniques employed in studying problems of learning are many and varied, as are the problems themselves; and at the same time there is an implied belief among psychologists that there is a single learning process. The experimentalist is content to attack any problem which presents itself by any suitable technique, while the theorist seeks to discover the process underlying any and all forms of learning. It is possible that the theorists are over-ambitious in their attempt, while at the same time it is possible that the experimental investigations of learning might be more significant if each could be given a setting in its more complete relationships. In view of these two possibilities, it seems worthwhile to examine the various types of learning to discover as far as possible their similarities and differences.

Whether or not there is a single process underlying learning, it is obvious that the end results from learning are varied. In the first place, learning may result in ceasing to do something as readily as in attaining new responses or greater perfection in old ones. In the second place, there are marked differences in the complexity of the attainments in the actions or ma-

[1]A Maller subtest.

terials learned. The action may be (*a*) a simple (single) response, (*b*) a stereotyped serial response, (*c*) a flexible set of actions by which the individual reaches a goal, or (*d*) a coordination of responses. On the basis of these two differences in what is learned, seven types of learning may be distinguished: (1) substitution of stimulus, (2) substitution of response, (3) negative adaptation, (4) serial habit acquired from serial stimulation, (5) serial habit acquired in trial and error, (6) flexible goal-reaching behavior, and (7) coordination. In our examination of these forms of learning we shall be interested in their similarities and differences and in the conditions under which they occur.

The first two types of learning are most clearly revealed in the conditioned-reflex studies and are very closely related. Substitution of stimulus (1) is the phase of conditioning most commonly emphasized and needs no comment here except to re-emphasize the fact that it is not so simple as it is sometimes made out to be. Substitution of responses (2) is a regularly appearing aspect of conditioning not so often considered. If one represents the stimulus-response relationships in the conventional fashion, both aspects of conditioning may be diagrammed as follows:



When one takes note of S_1-S_2-R_1, particularly of the new relation between S_2 and R_1, one speaks of substitution of stimulus; and when one is interested in S_2-R_2-R_1, one speaks of substitution of response. It involves the dissociation between a stimulus and its original response and the association of this stimulus with some other response. As a stimulus (S_2) gains effectiveness for response (R_1), it loses its effectiveness for its original response (R_2). Whether or not this double change in stimulus-effectiveness involves two sorts or only one sort of change in the nervous system cannot be said with certainty, but the studies of conditioning seem to point definitely to the fact that any change in stimulus-effectiveness is regularly a double change involving gain for one response and loss for another.

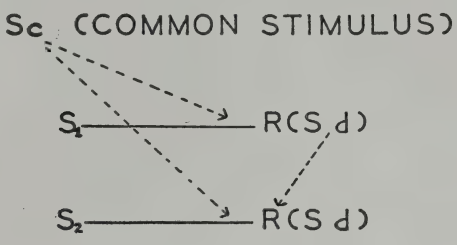
An apparent exception to this rule of double change occurs when a stimulus which once called forth a response ceases to elicit any response—occurs, if you please, in negative adaptation (3). Smith and Guthrie (3) give three outstanding conditions under which negative adaptation will take place: (*a*) if a stimulus is present in a weak intensity, (*b*) if it is present in conjunction with a strong incompatible or counteracting stimulus, or (*c*) if it is present when its response or the neural pathways to its response are fatigued. An examination of these conditions for something in common between them reveals this similarity: in all cases the stimulus is present and for one reason

or another is not followed by its regular response. And if it is not followed by its own response, it must be followed by other activities. In (*a*) some one of the multitude of other stimuli will determine the response, in (*b*) the counteracting stimulus will determine the response, and in (*c*) as in (*a*) some other stimulus will evoke a response. When negative adaptation occurs, a stimulus (S_2), instead of being regularly followed by a given response (R_1) dominant over its own as in substitution of response, is followed by numerous and varied responses (R_n). If one always read the moment the radio was turned on, the sound of the radio would soon be a contributory stimulus for reading instead of evoking its normal response of listening, i.e., substitution of response would occur. But if one carries on various activities —reading, eating, conversing, doing housework—while the radio is on, one becomes negatively adapted to it. Because the stimulus is followed by numerous activities, it gains only a slight, not a critical, effectiveness for each, and with these many slight gains goes the loss of effectiveness for its original response, just as in substitution of response a loss goes with a gain.

It is thus seen that the first three types of learning are closely related. The learning involves a change in stimulus effectiveness for given simple responses and the requirements for inducing the learning are simple, involving primarily the presentation of ineffective stimuli along with already effective ones. The *secondary* requirements have to do with controlling the conditions of the animal so that the stimulus designed to be effective will be effective, e.g., if food is to be employed as the effective or unconditioned stimulus for salivation, one must see to it that the animal is hungry and undistracted when offered food and that the food appeals to the animal.

When, however, one comes to a consideration of the other types of learning, complications begin to enter. Serial habits, for example, may be acquired in either of two widely different ways. They may appear as the result of the repetition of a sequence of effective stimuli or as the result of so-called trial-and-error learning. When they are acquired from repeated stimulation, they are closely related to the first three types of learning. In laboratory studies such acquisitions are best exemplified in the learning of a series of nonsense syllables. The training procedure is essentially the visual presentation of a syllable (S_1) with the subject instructed to pronounce it either aloud or subvocally (R_1), then the visual presentation of the second syllable (S_2) and its pronunciation (R_2), and so on through the list. After a number of repetitions (depending on the length of the series, etc.) the subject makes the responses in sequence when the first stimulus is presented or when requested to repeat the list. It seems evident that the chief difference from substitution of stimulus is that the stimulus substituted is the preceding response or stimuli (kinaesthetic, auditory, or whatever) derived from this response. This difference is readily expressed in the conventional diagram in which the broken line extends from R_1 or Sd (derived stimuli)

to R_2 instead of from S_1 to R_2. A second difference appears in the fact that the subject can repeat the list when requested. The first response at least must be set off by this request, and this means that R_1 must be conditioned to some preceding stimulus. This stimulus can well be the whole complex of conditions under which the list was learned. And if R_1 is linked to this stimulus-complex or common stimulus, it is to be expected that R_2 would be slightly less strongly linked with it, R_3 still less strongly, etc. Thus, the full diagram for a serial habit would be as follows:



The serial habit is not, then, a series of responses tied merely to each other but a series all tied to a common stimulus or situation. In its essential features it belongs with our first three types of learning in that it involves changes in stimulus-effectiveness for given simple responses, the changes being induced from the presentation of ineffective stimuli along with already effective ones.

The objections to considering serial habits of this first type (4) as belonging to the same category with simple conditioned responses derive from two sources. First, serial habits which have been built up in animals have been established through trial-and-error learning instead of through the use of the conditioned-reflex techniques. Second, the learning of a list of nonsense syllables is commonly classed as a voluntary activity while the conditioned-reflex studies have been made on involuntary or reflexive actions, and there is an *a priori* doubt that the same principles are applicable to both classes of response. But the author has just reported an experimental study employing the conditioned-reflex technique for a voluntary action and finding, it seems, that the conditioned-reflex principles are characteristic of voluntary as well as reflexive actions (2). After all, the difference between a voluntary and an involuntary action, as far as it need be made for experimental purposes, involves only a difference in the manner in which a given stimulus is made effective for a given response. When a stimulus is quite regularly followed by a given response, we speak of a reflexive or involuntary action; and when to get a stimulus to be regularly followed by a given response, we have to train or instruct our subject, we speak of a voluntary action. In short, the difference between those studies using nonsense syllables and those using unconditioned responses is in the secondary requirements for learning, not the primary ones. The secondary requirements have to do, as we have point-

ed out, with controlling the conditions in such a way that the stimulus designed to be effective will be effective. In both classes of action these secondary requirements must be given attention if learning is to occur, but the relative amount of attention required is greater for voluntary actions. This prominence of the secondary requirements should, however, not be allowed to conceal the fact that the primary requirements are essentially the same in both cases.

When we turn to serial habits acquired through trial-and-error learning (5), we come to a second method of learning, one which is common to our last three types of learning. In that learning so far discussed, the learner is encountering one set of stimuli which are already effective and which determine how he shall act at present, and he is also encountering a second set of stimuli which are ineffective but which determine what he shall learn, how he shall act in the future. He is being conditioned by the stimuli which he encounters, and his method of learning may well be termed the method of being conditioned. But in other cases commonly called trial-and-error learning, he encounters stimuli only in the course of his inner-initiated activity, and the way in which he will respond to these stimuli when encountered is to an important extent determined by his own inner set. Further, he is driven to activity until he encounters the particular stimulus for which he is set to react positively. The hungry rat in the maze exemplifies this method of learning. He is driven to activity by something within him, presumably his hunger or consciousness of hunger, and he continues until he finds food to which he is set to react positively.

There is little agreement as to the nature of the "drives" which initiate and sustain action and the "sets" which determine whether abient or adient responses will be made to the encountered stimuli. The writer has previously described what seems to him to be their probable nature (1). The primary drives, he has held, are physiological conditions or stimuli arising therefrom, and sets are neuro-muscular readinesses for given types of action conditioned to these internal stimuli, conditioned by virtue of the fact that the given action has regularly followed the particular internal condition. Neuro-muscular readinesses may also be conditioned to external stimuli, as for example the readiness to play golf when something associated with the playing of golf is seen or heard. These externally conditioned readinesses serve as secondary drives. In fact, it seems that the distinction between a drive and a set practically disappears upon close examination. Hunger per se is a drive only to heightened random activity, but hunger when conditioned to eating, when producing a readiness to eat is a drive toward food, or, if you please, is a set for eating. And further, if hunger has been regularly followed by travelling in a certain direction or to a certain locality, it becomes a conditioned stimulus to arouse a readiness for traveling in that way. Thus, a set or drive is seldom simple but rather involves several means-end-readinesses (Tolman) as well as an end-readiness. A readiness is a partially evoked action maintained through persistence of its stimulus,

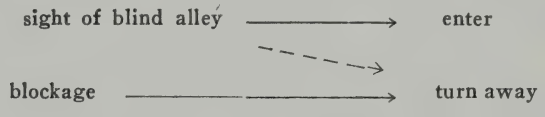
through the persistence of the neural gradient (Child), through circular arcs, or a combination of these. It is only partially evoked since it is a complex action whose component parts are conditionally linked to definite external stimuli (Tolman's discriminanda and manipulanda). We do not eat merely because we are hungry or merely because food is present, but rather when we are set to eat by virtue of hunger or an externally aroused readiness and when food is also present.

Whether or not the reader agrees with this analysis of the nature of drives and sets, he will hardly dispute the fact that both are something within the learner and that both are important in trial and error learning. But, however important they are for learning, they in themselves do not account for the learning. They are, in fact, to be classed with our secondary requirements for learning. In experiments employing the trial-and-error form of learning, the experimenter, instead of bringing the stimuli to the animal, controls the state of the animal so that it will go out to the various stimuli. Furthermore the stimuli which may be encountered by the animal are numerous, and the order in which they may be met may vary from trial to trial, and this contrasts sharply with the simple pair of stimuli presented to the animal in conditioning experiments. If one selects pairs of stimuli encountered by the animal in the course of its trial-and-error activity, the principles of conditioning are seen to hold. One of the pair will be effective for causing a given response in an animal with the particular "set" possessed by the animal in question, and the other stimulus of the pair will be ineffective but with repetition will take on effectiveness. For example, encountered blockage or pain stimulation will cause a hungry animal to turn away, while the sight of a blocked or grilled alley will cause turning away only after these visual stimuli have been encountered in conjunction with blockage or shock. Or again, encountered food will cause a hungry animal to move forward rapidly and unhesitatingly, while the sight of an open alley will cause such movement only after these visual stimuli have been encountered in conjunction with food. In other words, trial-and-error learning may be regarded as the establishment of a number of conditioned responses. Each conditioned stimulus may be regarded as a "sign" (Tolman) and each conditioned response as the development of an "expectation" or "readiness," but so also may the conditioned stimulus and response in simple conditioned-reflex experiments. And the whole of trial-and-error learning may be regarded as the development of "sign-gestalts" (still Tolman) and "means-end-fields," but this is only a cumbersome way of saying that it involves the development of a number or a complex of conditioned responses.

If this analysis of trial-and-error learning be at all correct, then the fundamental or primary requirements for learning are the same in all types and methods of learning, and the theorist is justified in his search for an essential process in learning. The primary requirement is that the learner encounter pairs of stimuli one of which is effective for a given response and the other of which becomes effective, or perhaps better that an ineffective

stimulus regularly precede a given response. In some cases (simple conditioning) these pairs are presented from the environment while the individual remains relatively stationary, while in other cases (trial and error) they are presented from the environment only as the individual actively explores that environment. Further, in the former cases the same pair is repeatedly presented under pretty much the same conditions, while in the latter cases a given pair may be encountered in one exploration and not in the next or it may be encountered at one point in one exploration and at a different point in the next.

It remains now for us to differentiate the three types of learning arising from trial and error. This task seems relatively simple. If the field of activity is so restricted that the animal can proceed to the terminal stimulus in only one way even though there may be culs-de-sac along that way, a stereotyped serial habit (5) will be formed. It will differ from serial habits established from repeated serial stimulation (4) only in the fact that in its development there has been a conditioned elimination of cul-de-sac entrances. Some stimuli which were encountered along the way and originally evoked adient responses come through conditioning to elicit abient responses, for example:



To each blind alley a conditioned abient response will be developed, and to each open pathway a conditioned unhesitating adient response will be formed. As these conditionings are established, the animal will repeatedly perform the same sequence of actions, and each action in the sequence will become linked to the stimuli derived from the preceding one.

If the terminal stimulus may be reached in either of two or more ways, adience will be developed for the stimuli encountered along each way, and the result will be a flexible, nonstereotyped goal-reaching habit, or, if you please, a set of alternative or equivalent habits (6). What is baffling about the fact that an animal can learn to attain a given end in any of several ways, the writer cannot understand. Each time an animal with a set for food reaches food, there is a conditioning of adience toward all stimuli encountered in reaching food, except toward those stimuli immediately paired with already effective abient stimuli. If there are two or more courses to the food and the animal follows each course, he will become adiently conditioned to each course.[1] There isn't anything to cause the development of

[1] The amount of adient conditioning toward any stimulus encountered along the way will depend upon its nearness in time to the reaching of the goal—Hull's goal gradient hypothesis. But the amount of conditioning for each abient response in no way that I can see depends upon its nearness to the goal. There are, however, reasons for expecting abient responses near the beginning to be developed first. Because of adient conditioning culs near

abience toward either course, but naturally if the course is so simple or the animal so highly intelligent that a course is learned in one trial, the animal might very well continue to use that course to the exclusion of others. It is solely when there is only one way to a goal or when the way of reaching that goal is learned in one trial that a stereotyped habit for reaching that goal should develop.

The development of coordinations (7) differs in no essential respect from the development of serial habits. If the terminal stimulus is to be reached only after a simultaneous rather than sequential combination of actions has been achieved, a coordination rather than a serial habit is developed. There is a learning to react to two or more stimuli at once, and at the same time there is a learning not to react or to react negatively to stimuli at first reacted to positively. These kinds of learning occur by virtue of stimulus pairs which are encountered in the trial-and-error performance.

A further word may be added with reference to the phrase terminal stimulus. It is the stimulus which the individual is set from the beginning to react to positively and which, when reacted to, removes the set, for when a readiness goes over into overt action the readiness no longer exists. The terminal stimulus is essential not only to the termination of activity but also to the development of adience toward the stimuli preceding and accompanying it. The encountering of food by the hungry animal not only terminates activity but also is essential to the establishment of positive reactions toward those stimuli encountered in reaching the food. Without pairing of stimuli there is no learning, and without the terminal stimulus there is no effective positive stimulus with which the various ineffective stimuli along the way may be paired. Without the terminal stimulus there would be only effective negative stimuli and in consequence only the development of conditioned abient responses. When it is said that for trial-and-error learning there must be a goal or a knowledge of results (which is the same thing), it is being said that if something positive is to be learned there must be a positive stimulus.

A final problem concerns the question of whether trial-and-error learning in man is comparable to this form of learning in animals. There is a general feeling among psychologists, and the writer concurs in this belief, that the difference is mainly or entirely in the number and kinds of sets underlying man's trial and error efforts. Man has many more sets (purposes, wishes, desires, wants, motives, incentives, or whatever), and of this number a large proportion are secondary rather than primary.

Summary

On the basis of what is learned, seven types of learning have been differentiated and each has been analyzed in terms of the changes in stimulus-effectiveness involved. (1) Substitution of stimulus, the development of an

---

the end will be eliminated early, and because of abient conditioning those near the beginning will be eliminated early.

effectiveness of an *S* for a given *R*, (2) substitution of response, the loss of effectiveness of an *S* for one *R* while an effectiveness for another *R* is gained, (3) negative adaptation, a loss of effectiveness of an *S* for one *R* while a less-than-critical effectiveness for many *R*'s is gained, (4) serial habits developed from serial stimulation, the gaining of effectiveness of a situation for many *R*'s in sequence and of response-derived stimuli from each *R* for the succeeding *R*, (5) stereotyped serial habits developed from trial and error, the same as 4 plus the loss of effectiveness of many elements of the situation for various *R*'s once exhibited in that situation, (6) alternative, flexible, or equivalent serial habits, the same as 5 except that there are two or more habits of reaching the same goal or that at some point in the habit there are two or more ways of proceeding, and in consequence the various *R*'s are not firmly linked together, and (7) coordinations, the same as 5 except that the parts of the habit are exhibited simultaneously or in overlapping fashion rather than in sequence.

Two methods or forms of learning have been described: (*a*) conditioning and (*b*) trial and error. The first four types of learning result from conditioning, and the last three from conditioning occurring in the course of trial and error. In both forms of learning the primary requirement for learning is the stimulation by an effective and an ineffective stimulus, or better, it is a sequence of ineffective stimulus and dominant response. In both forms the secondary requirements, those having to do with the control of the condition of the learner and his environment in such a way that the desired response will occur, are essential, but these requirements are more prominent in trial-and-error learning. In conditioning, the animal is relatively stationary and more or less subjected to the conditioning stimulation, while in trial and error he is active and encounters the conditioning stimulation by virtue of this activity. In conditioning, only one or a very few responses are being conditioned, while in trial and error many conditioned responses are being established more or less simultaneously.

This analysis contrasts trial-and-error learning with conditioning, not with the various types of learning. It reveals two kinds of changes essentially present in all types of learning and in both methods of learning: (*a*) the gain of effectiveness of a stimulus for a given *R*, and (*b*) the loss. It seems that as a stimulus gains effectiveness for one *R*, it loses effectiveness for other *R*'s, that both kinds of change are an expression of some single underlying process. The different types of learning may be regarded as results of the extent to which each of these changes occur and the number of stimuli for which such changes take place. Finally, this analysis has made a seemingly clear distinction between the primary and the secondary requirements for producing these changes.

REFERENCES

1. REXROAD, C. N. Goal-objects, purposes, and behavior. *Psychol Rev.*, 1933, **40**, 271-281.
2. ———. Reaction time and conditioning. *J. Exper. Psychol.* (In press.)
3. SMITH, S., & GUTHRIE, E. R. General psychology in terms of behavior. New York: Appleton, 1921. Pp. xii + 270.

Stephens College
Columbia, Missouri

THIRST AS AN ARBITRARY DRIVE

B. F. SKINNER*

In experimenting upon organic behavior it is often necessary to make an arbitrary selection of a basic drive. Hunger is frequently used. Its strength is relatively easily controlled through feeding and fasting, it can be graded through a wide range of strengths, similar strengths can be obtained repeatedly, and a given constant strength can be maintained through a considerable period of time even in the event of an occasional reinforcement. In some or all of these properties it possesses advantages over such drives as sex, escape from various kinds of negatively reinforcing stimuli (cold, electrical shock), general activity, and others. Thirst resembles hunger and apparently possesses one advantage of its own. Where hunger is in reality composed of many different specific hungers (salt, sugar, and so on), which make a formulation difficult, thirst (when freed of any hunger component) is presumably uniquely related to the ingestion of water. The experiments reported here test the resulting supposition that thirst might be supposed to be preferable as an arbitrary basic drive.

In a series of experiments reported elsewhere the behavior of a white rat has been studied with respect to the factors controlling the response to a lever, where the response has been reinforced with food. Eight of these experiments have been repeated upon small groups of rats in a preliminary survey of the same reflex based upon thirst, of which four will be briefly reported here. The original work and the technique are described in the references. The only modification in the present case is that the food magazine is replaced by a device which discharges a small amount of water upon a watch crystal in the experimental box, where it is

*Society of Fellows, Harvard University.

available to the rat. This device is a valved siphon composed of the following parts in order: (1) a reservoir of water, (2) a valve permitting a fast flow and operated by an electromagnet, (3) a branch line containing a small variable air space, (4) a constriction of the main line permitting a slow flow (say, one drop per second), and (5) an outlet 18 inches below the air space and reservoir. The amount of water delivered at each operation of the valve is determined by the air space. With the valve closed a negative pressure develops in this space, due to the weight of the water between it and the outlet. When the valve is momentarily opened, the pressure adjusts to the atmosphere, the flow at this moment through the constriction being negligible. During the subsequent few seconds a negative pressure is again developed and a measured amount of water discharged at the outlet.

Essentially maximal states of thirst can, of course, be produced by prolonged deprivation of water, or, more conveniently, by deprivation of both food and water and provision of dry food shortly before an experiment. The degrees of thirst used herein were in general considerably less than maximal and with few exceptions were obtained by allowing the rat to drink freely during a given hour each day and by experimenting just before that hour. Food was constantly available except during the experiments.

Conditioning. The original conditioning of the response to the lever can be accomplished by reinforcement with water (five cases). Evidence of a considerable amount of conditioning due to a single reinforcement is not always obtained because the effectiveness of the reinforcement may depend upon accidental aspects of the response selected, but some slight effect is usually observed and in many cases may be pronounced (1). The experiment upon this point reported in (2) was repeated upon four rats, of which one showed an even greater change than that previously reported with hunger. The record is shown in Figure 1*A*. During the first 33 minutes three responses occurred but were unreinforced. The fourth, at the arrow, is the only response in the history of the rat to be reinforced. It is followed by an extensive extinction curve, which is to be compared with Figure 1 in (2).

Extinction. This method of measuring the amount of conditioning shows incidentally the nature of the extinction curve obtained when the reflex is based upon thirst. Seven curves after considerable amounts of conditioning were also recorded. A typical curve after three successive periods during which all responses were reinforced is given in Figure 1*B* and may be compared with Figure 1 in (3).

Transfer from Hunger to Thirst. In several cases rats were taken from experiments involving hunger. A typical record from a group of seven is

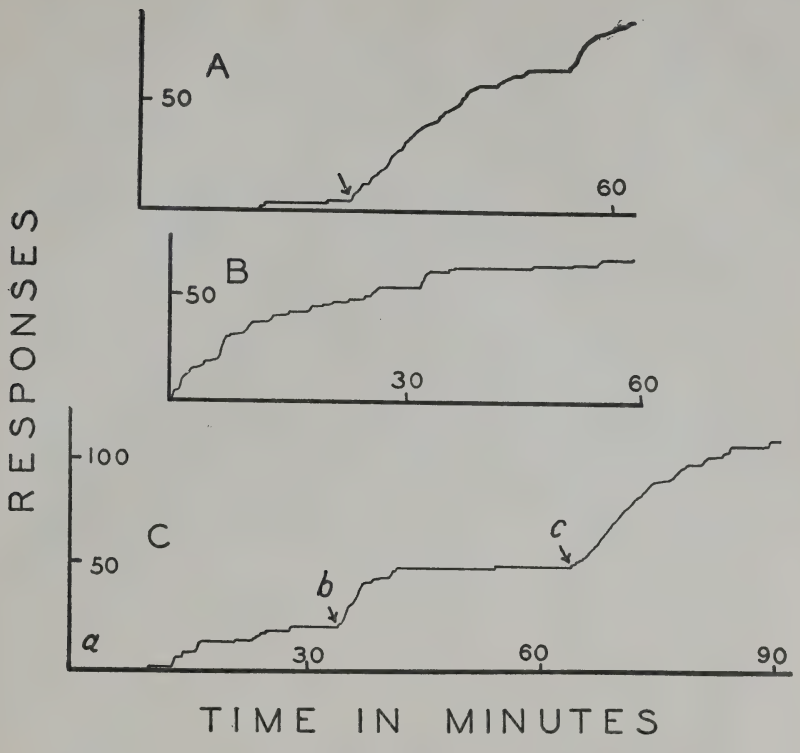


FIGURE 1

shown in Figure 1*C*. The rats in this group had been conditioned with food three months earlier and had developed various discriminations based upon hunger. They were then placed on a thirst regimen and trained to the sound of the magazine by reinforcement with water *in the absence of the lever*. In Figure 1*C* the rat was released at *a*. No water had been received by the rat in the presence of the lever prior to this experimental period. Food was constantly available (and in this experiment was present in the experimental box). The responses of the rat during the first 38 minutes were unreinforced. Since no conditioning with water had yet taken place, this initial responding may be regarded as a transfer of strength from one drive to the other. The rate shown in this figure is almost exactly the average for the seven rats (34 responses per hour). The early delay is characteristic of five records. The slight negative

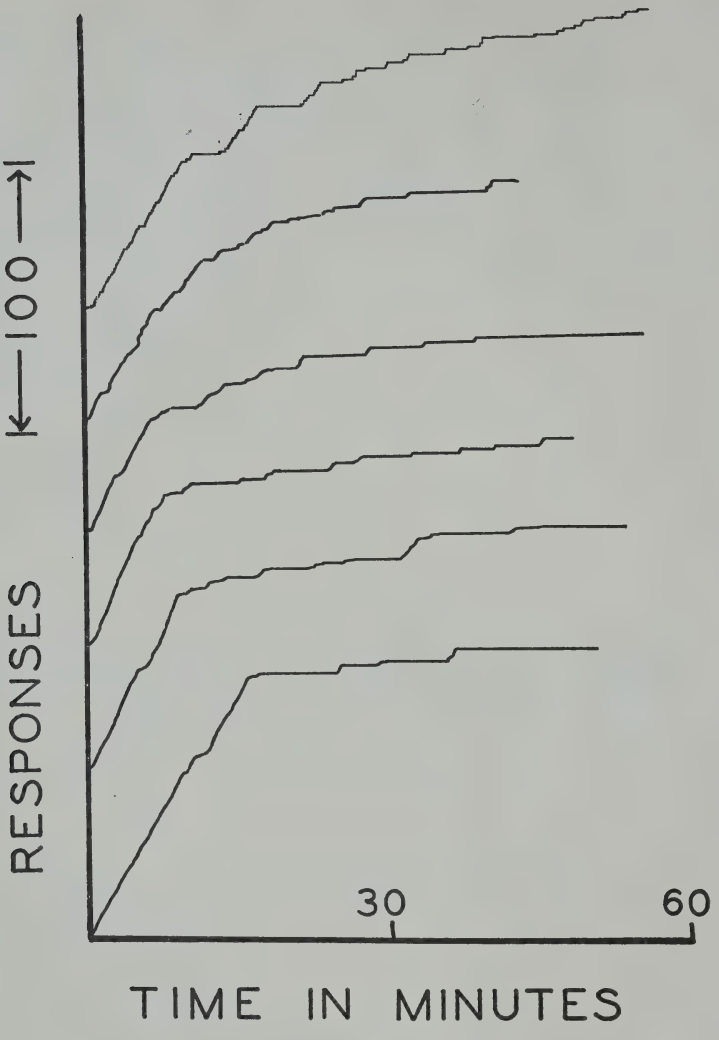


FIGURE 2

acceleration during this time is also characteristic and presumably indicates the extinction of the strength that has been borrowed from the other drive.

At *b* a single response was reinforced with water, and a subsequent extinction curve was obtained. A clearly observable effect of the reinforce-

ment of a single response appeared in all seven records. This part of the curve may be compared with Figures 1*A* and 1*B*. At *c* the reinforcement of every response was begun. A maximal rate of responding was developed and was followed by a gradual decline as the drive was weakened through drinking (see below). A marked positive acceleration may be noted at *c* as the newly conditioned reflex increases in strength. This is characteristic of all seven records and is, in general, to be observed whenever reconditioning follows extinction.

Change in Thirst as the Result of Drinking. A change in the strength of the response due to the ingestion of water may be shown simply by giving the rat free access to the lever and reinforcing every response. The resulting curves may be continuous, or they may show a single discontinuity. A series of records reaching from one extreme to the other is shown in Figure 2. The continuous curve is similar to (although not identical with) the typical curve obtained when the reflex is reinforced with food (4). No effect comparable with the discontinuous curve has been observed with hunger, although it was originally noted (5) that "a limiting factor brings each record to an end, either suddenly or after a short period of reduced frequency." The distinguishing characteristic of the present curve is the maintenance of a high rate up to the point of discontinuity.

The other experiments performed with water involved periodic reinforcement and gave essentially comparable results. But in spite of this successful repetition some reason to prefer hunger as a practical arbitrary drive appeared. The intensities of drive used in the two series were of the same order of magnitude, if we may judge from the extinction ratios, which give a convenient means of comparing the intensities of different drives (6). But the regularity of the behavior was noticeably less in the case of thirst. In particular, short periods of no responding, followed by compensatory increases in rate, occurred with greater frequency. It was also found to be difficult to produce the same strength of drive on successive days, although this may have been due solely to a failure to control the humidity of the living quarters.

A rough survey, therefore, indicates that thirst closely resembles hunger as a factor in the control of behavior but that for minor practical reasons it is not to be preferred as an arbitrary drive.

References

1. Skinner, B. F. On the rate of formation of a conditioned reflex. *J. Gen. Psychol.*, 1932, **7**, 274-286.
2. ———. "Resistance to extinction" in the process of conditioning. *J. Gen. Psychol.*, 1933, **9**, 420-429.

3. ————. On the rate of extinction of a conditioned reflex. *J. Gen. Psychol.*, 1933, **8**, 114-129.
4. ————. Drive and reflex strength: II. *J. Gen. Psychol.*, 1932, **6**, 38-48.
5. ————. On the conditions of elicitation of certain eating reflexes. *Proc. Nat. Acad. Sci.*, 1930, **16**, 433-438.
6. ————. Conditioning and extinction and their relation to the state of the drive. *J. Gen. Psychol.*, 1936, **14**, 296-317.

Harvard University
Cambridge, Massachusetts

SLEEP, WORK, AND FOOD HABITS IN THE TROPICS

RICHARD WELLINGTON HUSBAND

The effects of various environments upon human behavior are always interesting. The writer has recently completed a trip around the world in which he spent several months in the tropics, and during which he noticed several interesting modifications of living habits necessitated by the extreme and sustained heat.

I. SLEEP

The high degree of heat, complicated in some places by excessive humidity proves in the long run to be very debilitating. Continuous work throughout a whole day proves to be impossible. I was surprised to discover in Greece and Egypt, the first of the hot countries I visited, that all shops are closed during the early afternoon. A store may be open from 8 A.M. to 12:30 P.M., and from 4 to 7 P.M. Between these hours the workers go home and sleep about two hours. This does not represent an addition to the total amount of sleep, however. The people seemed to go to bed later and to get up earlier, on the average, than do Americans or Northern Europeans. They go to bed around midnight, yet are able to arise shortly after 6 A.M. The native quarter of a city in India is as busy at 7 as New York or London at 10 A.M.

These facts were of especial interest to the writer, as they furnished empirical corroboration to a bit of research completed just before departure on the trip (1). We had tested to see if sleep could be broken up into two four-hour installments, rather than taken in the eight-hour stretch we traditionally accept, and found that there was no measurable difference.

As to my personal reactions to the heat, I found that sleep during the afternoon eventually became necessary. I had never slept during the daytime, so attempted to go without it while in the tropics. But after a month I found myself becoming progressively more tired, and was forced to sleep, or lie down, for an hour or two in the afternoon. This seemed due both to heat exhaustion and to the heat cutting off sleep at both ends of the night.

II. Work

Work is carried on at a much slower pace in the hot climates. Even the routine habits of life, dressing, shaving, and household chores, are conducted in "slow motion" fashion. The prize example that I saw was two Kashmiri natives teaming up on one shovel, one manipulating it in the usual fashion, and the other assisting him by means of a string attached to the neck, where the handle joins the blade. Together, they were lifting about two pounds of loose dusty soil in the roadway! One is horrified at the extremely low wages paid native laborers, and at the numbers required to do a certain piece of work. But what they accomplish is worth no more. Granted they are less efficient than white workers would be; at the same time one would quickly suffer sunstroke if he attempted to play tennis, mow the lawn, or even walk down the street, at the pace he uses in more temperate climates. Even non-physical acts, like clerical tasks in a bank, seem to require much longer than one would expect. I heard several American tourists comment after seeing these things that they understood why our Southern States did things at a slower pace, and say that they would never again make fun of the "slack" Southerner.

III. Food

While the majority of equatorial countries I visited were British, and the Britisher likes to have things as at home, some dietary concessions are made. Very spicy dishes are served regularly, contrary to what one might suppose would be needed in hot climates. Curry-and-rice is eaten by the majority twice daily. They claim it serves to pep up the liver and other internal organs. One might also suspect that it may also serve to cover up a certain lack of freshness of the meats, or to stimulate a rather jaded appetite! The amount eaten is also surprisingly large, especially in contrast to the very light meals most of us in temperate climates eat during the hot summer months. It may well be that long-continued heat uses up so much energy that over a period of many months or years one must absorb greater quantities of food. We can afford to lose a few pounds during our short hot summer, but could not continue for years on end.

Just what this means in terms of metabolism is uncertain. Theoretically, one should use up less energy, as the share normally diverted into keeping up body temperature is not needed. Possibly the continuous perspiration drains out energy, even though muscular activity is not involved. It has been suggested to me that the residents of the tropics eat more out of habits formed during their earlier lives at home, but since they come from several nations, with different eating habits, and all exhibit this tendency, it would seem more of a genuine physiological phenomenon.

One cannot help noticing, also, that the white people in equatorial re-

gions consume tremendous quantities of alcohol. This may be partially due to monotony and loneliness. There are both positive and negative reasons given to account for this undeniable fact. They claim it is not harmful, as the heat enables one to oxidize, or sweat out, what appears to an outsider very damaging amounts. As a benefit (or rationalization), it is said, in common with spicy foods, to stimulate the digestion.

REFERENCE

1. HUSBAND, R. W. The comparative value of continuous versus interrupted sleep. *J. Exper. Psychol.*, 1935, **18**, 792-796.

University of Wisconsin
Madison, Wisconsin

PSYCHOPHYSICS AND MESCAL INTOXICATION

THEODORE KARWOSKI

The study of mescal intoxication has given us a most interesting literature. The descriptive terms applied to mescal such as the "Divine Plant," the "Diabolical Root," an "artificial Paradise," etc., connote a state of mind which is so extraordinarily different from that of the normal that nonmundane terms are used to describe it. Many accounts of the wonders of mescal in producing colors, sounds, kinaesthesis, hallucinations, distortions of space and time, feelings of grandeur, etc., have been reiterated in one report after another. Mescal has been proposed as an instrument of war since it induces insomnia and indefatigability lasting over 24 hours. It has been offered as an aid in the field of medicine in the form of a sedative for nervous subjects and as a means of establishing rapport with the doctor in a phychoanalytic seance. To students of psychology and psychiatry it offers extraordinary conditions for establishing vivid imagery and presents by actual experience some glimpse of the abnormal mind in the exaggerations and the fancies obtained under mescal intoxication. To the sexologist an easy way of obliterating temporarily the genital response is offered since mescal is a powerful an-aphrodisiac. In fact it has already been pointed out that the ceremonial ritual of Indians who partake of mescal is peculiarly devoid of sexual symbolism and implications. My own experience confirms the an-aphrodisiac properties of mescal, but the fact that under its influence I found my imagination turning to erotic situations, although temporarily impotent, is an illustration of the persistence of conditioning that offers an interesting suggestion with reference to the

extirpation experiments reported in the controversy over the James-Lange theory of emotions.

Unfortunately an analysis of the literature on mescal intoxication reveals glaring differences between individuals, so that all the generalizations made are contradicted by exceptions to the rule. No studies involving careful measurements have been reported; at least, the quantitative data from such studies has not been presented, though sporadic statements of experimental tests have been made. Much of the work done in this field has been rather haphazard, since the main interest of the experimenter has been to take enough mescal to produce visions. In general this work has established the presumption that mescal affects particularly the occipital lobes of the brain, thus accounting for the subjective color effects experienced. The effects on the other senses have not received extensive attention, but there is abundant evidence that they also are modified. The question whether these effects are primarily central or peripheral has never been crucially tested, although Fernberger attempted to do so. The most obvious lacuna in the literature is the lack of psychophysical data. Nevertheless, generalizations have been made in terms of the predominance of certain kinds of data which have been found to recur in separate investigations. The most common of such generalizations is that mescal lowers the threshold for sensory experience. These claims are based on casual observations that lights seem brighter and sounds seem louder. But there are exceptions; and ample evidence could be assembled to indicate that the threshold is increased, particularly in the case of smell and touch, since anosmia and anaesthesia are frequently referred to. That individual differences are striking under mescal intoxication is well attested but inasmuch as little attempt has been made to control the amount of the dose of mescal in terms of the tolerance of the individual we cannot be certain that differences between individuals are not in large measure a function of the degree of intoxication as determined by the sensitivity of the subject for Peyote. For this reason, and because of the little progress that has been made so far in quantitative studies of mescal intoxication, it seems worth while to describe some recent experiences in experimenting with mescal, and to offer suggestions for improving conditions of experimentation.

Experimental Procedure and Results

The special conditions imposed by mescal in psychophysical investigations require a reliable method of determining rapidly the threshold for light while Peyote photisms are still present. A procedure was devised so that the general and specific color thresholds could be obtained within an experimental session of two hours. By using a spectral light obtained with large slit openings, a light having a dominant wave-length of 710 mu

and a narrow photochromatic interval was established. By means of a neutral tint wedge, which could be moved in front of the exit slit, intensity variation was easily controlled. It was found that both the general and the specific color thresholds were confined within a range of 12 mm. on a Wratten neutral tint wedge having a transmission range of 1 to 1000. This narrow range of 12 mm. was blocked off into 12 steps of 1 mm. to serve as step intervals for psychophysical measurements. Actually the method of constant stimuli was used. The general and specific color thresholds were both obtained in one experimental session. An artificial pupil of 2 mm. square was used because it is well known that Peyote dilates the pupil.[1]

The author and one other subject participated in the experiment. They shall henceforth be referred to as *K* and *E*. *E* is predominantly the visual type; *K* has poor visual imagery with kinaesthesis predominating. Thresholds were carefully determined for each subject on the day preceding mescal intoxication. The limens were based on 40 determinations per step in the constant stimuli series. With *K* two sets of data were obtained, one in the morning and one in the evening of the same day, in order to determine the reliability of the thresholds. The morning and evening limens proved to be practically identical.

In the mescal experiments both subjects took 18 grams of the button prepared in capsules. The drug was consumed in the course of three hours by taking six grams per hour. *E* vomited about six hours after taking the first dose; mescal photisms made their appearance soon after. *K* did not vomit but he felt nauseated and resorted to bed for relief. *K*'s photisms began to appear about seven hours after the first dose. They were not so vivid or so persistent as in the case of *E*, but for one who rarely experiences visual imagery, the rich color transformations, though miniature in size, were quite an engaging spectacle. *K* obtained photisms only under conditions of quiet, and preferably with closed eyes. *E* apparently was more profoundly obsessed with color presentations. They appeared to him at all times no matter what activity he was engaged in. *K* was more profoundly affected kinaesthetically.[2] One fact of correspondence

[1]For further details in the apparatus and procedure consult forthcoming article (5). The author is indebted to Dr. S. W. Fernberger and Dr. Mason Crook for his supply of Peyote (*Anahalonium lewinii*).

[2]Since very little has been reported on the effects of mescal on kinaesthesis, the following observations may be of interest. *K* enjoyed his kinaesthetic experiences. Contacts of all kinds were very pleasant. While reclining in bed he found himself stretching his legs and curling his feet. The pleasant sensation was accompanied by a perception of the pattern formed by the various muscles contracted. Probably apprehension is a better word, since the pattern was not actually seen though vividly felt. In this respect *K*'s experience was similar to that of Fernberger. However, after the effect of mescal wore off, contracting the various muscle groups,

between *K* and *E* was noted. Both tended to see patches of spectral colors in the field of vision under conditions of simple introspection. *E* reported weird and fanciful picture colors as the principle mental content while trying to fall asleep; *K* experienced color and sound presentations. Both *E* and *K* spent a sleepless but rather enjoyable night. Photisms persisted all night and were still present in less distinct form most of the morning.

Threshold determinations under the influence of mescal were undertaken with *E* an hour after the photisms were first reported. Immediately difficulties were encountered. It was soon apparent that *E* was confusing his photisms with the circular stimulus light. Furthermore, it seemed to be impossible to obtain a psychometric function with 1-mm. step intervals on the neutral tint wedge. The step intervals were increased to 2 mm. *E* still seemed to be quite bewildered, and his sense of orientation was so disturbed by the fleeting photisms that he did not know whether he was looking directly forward or not, even though his position was fixed by means of a biting board containing the dental impression of his teeth, and his eye was placed against an eyepiece containing an artificial pupil which furnished kinaesthetic cues of the position of the eye. Another change in the conditions was introduced; a fixation point was put in the system to locate the stimulus field. Even so, *E* found that colored photisms entered and left his field of vision so frequently that he had difficulty in establishing a criterion of the objectivity of the light. When he believed that he saw the stimulus light it changed colors so rapidly that he could not be certain that the color was not entirely imaginary. Finally the following criterion was worked out by the subject himself. The judgment was based on the first light seen on the field of stimulation, immediately after the ready signal was given. However, confusion still remained; one third of the total judgments were false responses to no stimulus. Nevertheless 25 judgments per step were obtained. The statistical results show that the limens for

---

though still pleasant, became painful if attempted with the same intensity as when under Peyote intoxication. In fact, *K* noticed that his hands were sore from the excessive wringing and squeezing which was indulged in because of the pleasure thus experienced. This raises the question as to the cause of the vividness of kinaesthesis. Apparently, under the influence of mescal, kinaesthesis is vivid and pleasurable, because, the sense of pain being dulled, kinaesthesis can be experienced at intensities ordinarily accompanied by pain. This selective action of mescal on cutaneous sensibility should be of interest to students of skin and muscle sensitivity. I have become particularly interested in my cutaneous sensibility since this experiment was done, and have noticed that the only time I can experience anything like the conditions described above is when I am especially fatigued. This would indicate that mescal is a depressant on the peripheral nervous system, which conforms with the findings in the studies of animals such as the guinea-pig and the cat.

the general threshold for light do not differ significantly between the normal and mescal series, but the limen for the specific color threshold was significantly lowered for the mescal series.

Subject *K* was not particularly troubled by photisms under actual conditions of experimentation. Observational activity seemed to inhibit them though they were likely to be present at any time; and strangely enough they tended to resemble the stimulus light in hue and shape. *K* was slow in responding, as was also *E*, and thus the rhythmical condition of stimulus and response obtained in the normal series was not maintained under mescal intoxication. *K*'s state of mind was one of happy indifference, with no zest in the activity of observing, in spite of the fact that he was seriously interested in the results of the experiment. The results obtained indicate a slight decrease in the sensitivity for both the general and specific color.

The results are difficult to accept on their face value. The changes introduced in the mescal series for each observer, especially that of making the steps in the constant stimuli process twice as coarse as those of the normal series, indicate that the subject's sensitivity was decreased rather than increased. The slower responses on the part of both subjects is another indication of the narcotic effect. Furthermore, it is unlikely, though not impossible, that the sensitivity for the specific color should be lowered, but not for general light as reported by *E*. Confusion between subjective and stimulus color may account for the lower limen obtained for red light. On the other hand, it is difficult to attribute *K*'s reduced sensibility entirely to the sensory mechanism of vision, because the state of pleasant abulia in which *K* found himself under the influence of mescal probably introduced changes of a more general kind.

The Implications of the Experiment on Methodology for Experimenting with Mescal

The difficulties involved in these experiments may be approached from the standpoint that mescal intoxication is equivalent to comparing an abnormal with a normal condition of the organism. If the pathological effects of the drug were circumscribed to the eye, then it could be assumed, for practical purposes, that only the retinal cerebral arcs were different in the normal and intoxicated condition. However, it is apparent that the whole system was so changed in these experiments that it would be difficult to interpret any change in the threshold, particularly if these changes are of small magnitude. This situation prevails in varying degrees in all psychophysical experiments, since the total organismic set or total response configuration is different with changed conditions of stimulation. This is particularly so in experiments demanding a division of the attention, or in experiments involving difficult judgments. Recently, experiments in the field of intersensory interaction have been reported in which the threshold of

vision was observed to change slightly when the visual stimulus was presented simultaneously with an auxiliary stimulus, sound for instance (4). This type of experimental situation, particularly, is subject to influences beyond the particular nervous mechanisms proposed for investigation. Consequently, the results, often equivocal in terms of consistency between different experimenters, have been attributed to complex states, such as attention, configurational properties of the visual field, cortical diffusion, etc., without precise information of the mechanisms involved (4). In the case of drugs the effects on the total organism are much more profound and it is much more difficult to circumscribe the experiment to some particular segment of the nervous system. Consequently, the question of methodology for psychological experimentation on the effects of drugs needs careful consideration.

The first principle I should propose is the principle of the *minimal stimulus.* Everything else being equal, experiments with mescal should be done with the minimal amount of the drug. In this way the widespread effects on the organism may be reduced and the investigation may, thus, be more definitely confined to a more limited part of the nervous system. The definition of the minimal amount is impossible to determine in the case of most drugs, since it varies among individuals and changes with addiction. However, in the case of mescal, the minimum may be defined as the quantity which just produces visual photisms, if the investigation concerns vision; or the amount which just induces subjective sounds, if audition is the subject of inquiry, etc. This principle is particularly important in the case of mescal because the effects on the general system are so widespread.

The second principle is a consequence of the principle of minimal stimulation. The amount of mescal which is necessary for different persons to just produce photisms may be regarded as equal since the effects are similar, even though one person may require more drug than another. Since these effects are phenomena such as photisms, subjective sounds, heightened kinaesthesis which can be observed, some precision may be attained in determining the amount of mescal different subjects should be given to attain threshold stimulation. This principle may be called the principle of *equi-effective stimuli.* It follows from this definition that mescal doses may be varied quantitatively for a group of subjects by taking multiples of the amount just necessary to produce certain phenomenological effects for each subject. Thus, if *A* requires twice as much mescal as *B,* the stimuli, theoretically, would still be equal, in terms of the threshold effects, if each were given a fractional part or multiple of the original dose.

In actual application these principles may not be adequate in individual cases, because the difference in the amount of drug which different people may require may be quite great. Thus *A* may require five times as much drug as *B.* Now a person who requires so much more drug than the aver-

age is not a good subject; with such a person there is no way in which powerful effects on the total organism can be avoided. Hence another guiding principle is necessary—the principle of *selected subjects.* It is possible to find people who require only small doses of mescal to produce photisms, and the experimenter should prefer to work with such subjects, unless for special reasons he is interested in extreme intoxication.

The literature on mescal is still very meager. It is difficult to obtain specific information about the effects of mescal to illustrate the application of the methodology here proposed. For this purpose facts on three points are necessary: (1) the effects from slight amounts of mescal, (2) the latent time of appearance of photisms, and (3) their duration. Data on the first question is particularly hard to obtain because most, if not all, experimenters have preferred large doses in order to be certain to obtain typical mescal effects. However, Kelly (6) reports that four of his five subjects obtained slight photisms as a result of five grams of mescal within an hour. Since these subjects were given another dose at the hour, information concerning the duration and vividness of mescal visions obtained from five grams is not available. Unfortunately, Kelly's reports are not confirmed by other investigators. Fernberger (2) obtained photisms from eight of nine subjects, and, though his subjects took heavy doses, the latent periods for visions ranged from 2 1/3 to 5 hours. In the experiment here reported the latent period seemed to be more than six hours for both observers. These discrepancies may be due to the method of determining the initial phases of mescal visions. Fernberger's subjects met at a Peyote seance, conducted ceremoniously after the manner of the Indians. It is doubtful that such surroundings presented a favorable condition in which to detect the first photisms. In the above experiments, frankly, no attempt was made to determine carefully the latent period. Kelly, on the other hand, had to assemble his subjects in a dark room at hourly intervals as part of the procedure of the experiment. This happy circumstance provided good conditions in which to note the initial phases of mescal intoxication. For this reason I am inclined to accept Kelly's data as typical. Concerning the duration of mescal photisms there is more agreement. Fernberger's (2) subjects ranged from two to 20 hours; Kelly's subjects reported photisms after 15 hours, and in my experiments both subjects retained visions at least 12 hours. On the basis of this analysis of the meager data available, it seems reasonable to expect that a sufficient percentage of the population may obtain photisms from five grams of mescal within a few hours after injection,[3] and enduring sufficiently long to permit experimentation.

[3]Prentiss and Morgan (8) describe the effects of small doses of mescal on human subjects. They find that a dose from one-half to one gram is sufficient for therapeutic purposes. Such doses are recommended for general nervousness, nervous headache, nervous irritative cough, abdominal

It should be noted that with five grams of mescal the most unpleasant symptoms are eliminated, thus permitting repetition of experiments without much discomfort to the observers.[4]

On the basis of the above analysis of the quantity and time relations involved in producing mescal visions, it would seem that the methodological principles here proposed can be utilized successfully. For precise determination of the minimal stimulus the experimenter may administer a subminimal amount of mescal to begin with, and follow this dose with definite increments of the original dose at definite intervals until mescal photisms are reported. At the present stage of our knowledge the details of this procedure, concerning the size of the original dose, the magnitude of the succeeding doses, and the time spacing, must be worked out by trial and error. For practical purposes, an investigator who wishes to avoid such an elaborate preliminary preparation may utilize a simpler procedure that probably is not very much different from the plan here presented. Since no one has reported mescal visions with less than five grams of mescal, this quantity may be arbitrarily considered to be the minimal stimulus. By giving a rather larger group of people five grams of mescal, a sufficient number of individuals could be found who obtain photisms within certain time limits. Thus a measure of stimulation, though not as precise as we should like, is nevertheless available. In this sense mescal has an advantage over an extensively studied drug such as alcohol, which does not lend itself to being equated in terms of the psychological effects produced.

There are many complications other than those considered in this paper which enter into the general problem of experimenting with mescal. There is no doubt a correlation between quantity and the quality of the nervous reaction. In the case of alcohol, for instance, slight doses are suspected of facilitating nervous function; large doses of inhibiting function (1). This type of problem and others of a similar nature, in which large quantities of drug are necessary, may be the kind of problem in which the controls proposed here may not be serviceable. Nevertheless, in a field of investiga-

---

pain due to colic or griping of the intestines, hysterical manifestations, hypochondria, melancholia, and allied conditions. Thus slight amounts of mescal are generally beneficial to the human organism. The same authors state that in some cases doses large enough to produce intoxication and color visions (three buttons) have been found to be helpful in treatment of neurasthenia. There is little if any information on the effects of intermediate amounts of mescal (two to five grams).

[4]This analysis is made for visual symptoms only because few observers have been especially concerned with other modalities. Hence it would be hazardous to make any comparison between modalities though such specific information is desirable for methodological purposes.

tion in which so little has been done at all accurately there is much to be achieved within the limits proposed in this paper.[5]

References

1. Dodge, R., & Benedict, F. G. Psychological effects of alcohol. *Carnegie Instit. Washington,* 1915, **232**, 250.
2. Fernberger, S. W. Further observations on Peyote intoxication. *J. Abn. & Soc. Psychol.,* 1932, **26**, 367-378.
3. Grace, G. S. The action of mescaline and some related compounds. *J. Phar. & Exper. Therapeut.,* 1934, **50**, 359-372.
4. Hartmann, G. W. Changes in visual acuity through simultaneous stimulation of other sense organs. *J. Exper. Psychol.,* 1933, **16**, 393-407.
5. Karwoski, T. An investigation of the effect of auditory stimulation on visual sensitivity. *J. Exper. Psychol.* (In press.)
6. Kelly, E. L. Individual differences in the effects of mescal. *J. Gen. Psychol.,* 1933, **9**, 462-472.
7. Knauer, A., & Maloney, W. J. A preliminary note on the psychic action of mescaline with special reference to the mechanism of visual hallucination. *J. Nerv. & Ment. Dis.,* 1913, **40**, 425-436.
8. Prentiss, D. W., & Morgan, F. P. Mescal buttons. *Med. Rec.,* 1896, **50**, 258-268.

Dartmouth College
Hanover, New Hampshire

---

[5]Mescal buttons have been chemically analyzed into four alkaloids. Of those mescaline is said to be entirely responsible for the visual effects. Experiments with animals, however, indicate that the other alkaloids have effects similar to that of strychnine. Mescaline on the other hand, acts as a depressant on animals. See Grace (3). Although no effects on humans have been found except by mescaline, there is some chance that effects are present but not ordinarily detected. This may in part explain the strange individual differences as the result of mescal intoxication. For this reason it may be more defensible to use pure mescaline. Knauer and Maloney report results of mescaline injection (7).

APPARATUS

AN INEXPENSIVE COMPLICATION DEVICE

REX MADISON COLLIER

The complication device described in this article was constructed by the writer in the Psychological Laboratories of Northwestern University. Three reasons seemed to justify publishing a description of the apparatus: (1) the device can be used conveniently and successfully by the beginning classes, (2) it is inexpensive, and, (3) since the apparatus is inexpensive, students of the beginning classes may become acquainted with a type of psychological phenomena hitherto omitted from most of the manuals.

The accompanying photographs show the principal details of construction. Figure 1 shows the 12-inch clock dial and hour hand which may be purchased at a jewelry establishment. The motor is an induction phonograph motor for A.C., 110-volt current. It may be purchased from the Motor Player Corporation, Chicago, Illinois. The friction speed control allows a range in speed variation from 0 to 120 r.p.m. It was found necessary to alter slightly the motor as it is used in this device. The shaft which ordinarily drives the phonograph turntable must be removed and replaced after rotating it through an arc of 180 degrees. To replace the worm gear bearing, new screw holes must be drilled and tapped. When this change is made, the end of the shaft clears the framework of the motor to a distance of five inches. This distance is sufficient to allow the shaft to penetrate the pressboard which supports the dial, and to allow room for the accessory devices behind the dial necessary to regulate the moment at which the bell rings.

Figure 2 shows the position of this shaft after the above change has been made. The photographs indicate the construction details of the frame upon which the motor is mounted.

Figure 3 is included for the purpose of showing the method of controlling the moment at which the bell is rung during the sweep of the pointer. From a coil spring of No. 30 wire, ¼ inch in diameter, a piece is cut with a length of four or five turns. After a tip of about ¼ inch has been bent out and straightened, the base of this small coil is soldered to a pfanstock clip as shown in Figure 3. The coil so mounted acts as base for the projecting tip. This makes a flexible contact past which the contact pointer may move with practically no sound in addition to the bell and without being jarred out of alignment with the observation pointer. It is necessary that both pointers be kept in alignment if the bell is to be rung at a predetermined point. The pfanstock clip is mounted on a piece of No. 12

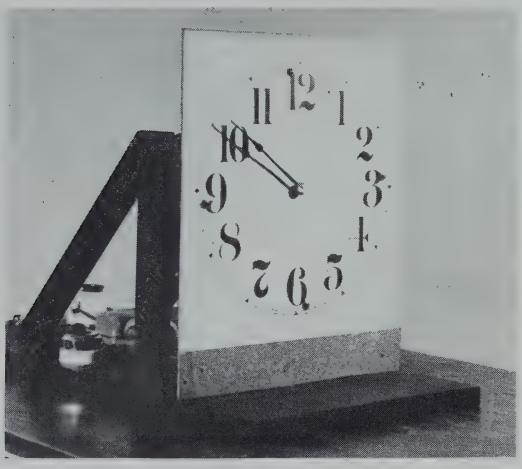


Figure 1

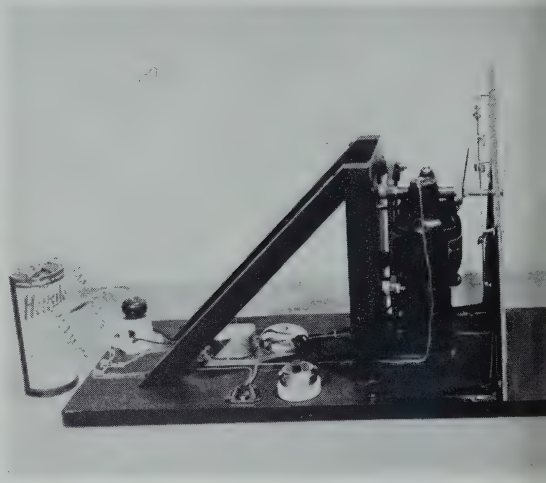


Figure 2

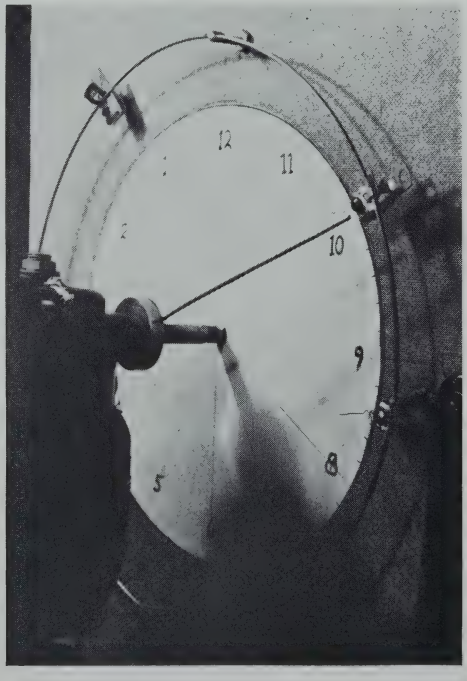


Figure 3

copper wire. Four such clips, one for each quadrant, were used. The posts upon which the No. 12 wire was mounted were displaced about five degrees to one side of the hour position. This reserves the possibility of using any of the hour positions for contact points. A clip which is not being used to make the bell circuit may be rotated on the wire so that the contact pointer does not strike it.

Dimensions. Base board—1′ x 11½″ x 23″. Motor support, upright piece—1″ x 1″ x 10″; diagonal brace—1″ x 1″ x 14½″. Wall-board backing for clock dial—11½″ x 14½″ x ⅛″.

Circuits. The motor itself is operated directly from the 110-volt A.C. current. A switch may be introduced to facilitate motor control. Figure 2 shows such a switch. The bell circuit is a simple series circuit with one lead passing through a push button and grounded directly to the motor. The other lead is connected to one of the posts supporting the wire upon which the pfanstock clip is mounted.

Northwestern University
Evanston, Illinois

BOOKS

GREGORY H. S. RAZRAN. *Conditioned Responses: An Experimental Study and a Theoretical Analysis. Arch. Psychol.*, 1935, **191**, 1-124.

In this monograph Razran described a series of experiments in which he conditioned the salivary response of human subjects to various stimuli, and under various instructions. The technique he devised to measure secretion was as simple as it was effective. An absorbent dental cotton roll was placed in the mouth and allowed to collect saliva for one minute. The amount of secretion was determined by weighing the cotton before and after insertion. The magnitudes of the conditioned responses were usually presented in terms of the secretion during "control" insertions of cotton, to eliminate the effects of the normal flow of saliva. In all, 37 subjects were used, yielding a total of 800 experimental hours. The results are given elaborate statistical treatment and are presented in 66 figures and tables.

Nearly all of the subjects showed some signs of conditioning, but uninstructed subjects were often sporadic and variable. The unconditioned stimuli employed may be listed in order of decreasing effectiveness as follows: eating pretzels, observing the eating of pretzels, keeping a mint on the tongue, thinking of eating steak, and thinking of eating pretzels. As conditioned stimuli a metronome, nonsense syllables, words, and digits were used. They varied somewhat in their effectiveness depending on the unconditioned stimuli with which they were coupled. Razran refers this variation to a "ratio principle." Social facilitation of conditioning occurred if two subjects were employed simultaneously, and there was a small additional gain when a third and a fourth subject were added to the experimental group.

To the reviewer, the most interesting results were those that had to do with the attitudes of the subjects. Razran found that conditioning (and extinction) did not proceed in a regular fashion, but that there was no relationship between the length of training and the magnitude of the secretion. After a few paired stimulations, one subject in a given experiment gave consistently large secretions, another almost none, and a third actually decreased his resting flow of saliva in response to the conditioned stimulus. On the basis of the distribution of their responses these subjects were classified as positive, indifferent, and negative types, respectively. A given subject might be classified in a different type for different sets of conditions. For example, three of the subjects were positive when in a group, but indifferent or negative when alone. Further, a subject might show a bimodal distribution of responses, even in a single experiment. Such a

subject would be a bitype. Three monotypes, three bitypes, and occasional tritypes were found. The form of the distribution was determined through the calculation of β constants. Razran assumed that the type of response a subject yielded was to be attributed to his attitude and attempted to determine the attitude by instructions. To a large extent he was successful. When subjects were instructed to form associations between the conditioned and unconditioned stimuli, very substantial and stable responses were obtained. Instruction such as "ZIG DAH means saliva in Benga, an African language," were less effective than "positive volition," but more so than no instructions. On the other hand, instructions not to form certain associations usually led to zero or negative conditioning. Particularly interesting were two subjects who reacted negatively to instructions.

To explain his results, Razran assumes that a conditioned response in an adult human subject is dependent on two factors: (1) a quantitative physiological factor, depending on the number of repetitions and the "linkage strength" of the two stimuli, and (2) a psychological, qualitative, "yes-no" factor. In experiments such as he has reported, the second factor is more important, explaining the conditioning types, the absence of regular learning curves, and the failure to show experimental extinction. Razran traces these two factors through Phylogeny and Ontogeny. The quantitative factor describes the conditioning of animals. But when the human child reaches the age of three to five years, there is a reversal in the trend toward rapid conditioning that is found as we come up the scale. This reversal is attributed to the emergence of the qualitative factor. The past experience of the organism, through symbols, builds up a "choosing and rejecting system." Razran believes that the action and organization of this system proceed ultimately according to the same principles that govern simple conditioning, that it is *"mechanical and summative in structure but emergent and supra-summative in function."* These attitudes may be the ultimate building blocks of the attitudes and traits found in personality analysis. At the end of his monograph, Razran elaborates this view into a theory of Mind.

In the reviewer's opinion, Razran has made an extremely important contribution in his clear-cut demonstration of a factor that accounts for the differences between simple conditioning in animals and man, on the one hand, and modification of "conscious or verbalized or voluntary reactions," on the other. The recognition of such a factor has been delayed by a somewhat naïve objectivism that neglected the importance of the language mechanism. The reviewer (and probably many others who have attempted to establish conditioned reflexes in man) has long been disturbed by the fact that a subject, unlike a dog, must be given instructions when he is placed in a conditioning situation. If such instructions are omitted, the subject is forced to give them to himself. During the last four years several experimenters have suggested that the subject's attitude may mask con-

ditioning, but none has analyzed the matter as clearly as has Razran. How far one wishes to follow him in his description of the development of Mind is a matter of theoretical tendencies, but it would seem to the reviewer that one must recognize the differences found at a *gross* level between conditioning in man and lower animals. The recognition of these differences may pave the way to their ultimate reduction to simpler processes.

In passing, it seems unfortunate that this work was published in monograph form. Limitation of space made necessary a curtailment of discussion in some places where Razran's knowledge of the Russian literature would have been most valuable.

HAROLD SCHLOSBERG

Brown University
Providence, Rhode Island

BOOKS RECEIVED

BLANTON, A. W. The child of the Texas one-teacher school. *Univ. Texas Bull.*, 1936, No. 3613. (*Bur. Res. Soc. Sci.*, 1936, No. 17.) Pp. 1+111. Gratis.

BOURDON, B. La perception. *Nouveau Traité de Psychologie,* 1936, **5**, 1-82.

BROWN, J. F. Psychology and the social order. New York & London: McGraw-Hill, 1936. Pp. xiv+529. $3.50.

COOKE, D. H. Minimum essentials of statistics. New York: Macmillan, 1936. Pp. xix+271. $2.75.

DELACROIX, H. La croyance: la psychologie de la raison. *Nouveau Traité de Psychologie,* 1936, **5**, 185-304. 20 francs.

————. Le temps et les souvenirs: le rève et la rêverie. *Nouveau Traité de Psychologie,* 1936, **5**, 305-404. 20 francs.

————. Les opérations intellectuelles. *Nouveau Traité de Psychologie,* 1936, **5**, 83-184.

DETWILER, S. R. Neuroembryology. New York: Macmillan, 1936. Pp. x+218. $3.75.

DUNLAP, K. Elements of psychology. St. Louis: Mosby, 1936. Pp. 499. $3.00.

FRYER, D., & HENRY, E. R. An outline of general psychology. New York: Barnes & Noble, 1936. Pp. 229+xliii. $.75.

GALLI, E. L'Estetica e i Suoi Problemi. Napoli: Alfredo, 1936. Pp. viii+391. 1.30.

GARRISON, S. C., & GARRISON, K. C. Fundamentals of psychology in secondary education. New York: Prentice-Hall, 1936. Pp. xx+599. $2.80.

GURNEE, H. Elements of social psychology. New York: Farrar & Rinehart, 1936. Pp. xi+467. $2.50.

HIGGINSON, G. DEV. Psychology. New York: Macmillan, 1936. Pp. xiii+646. $2.90.

HITSCHMANN, E., & BERGLER, E. Frigidity in women. Its characteristics and treatment. Washington & New York: Nerv. and Ment. Dis. Publ. Co., 1936. Pp. iii+76. $2.00.

JANET, P. L'intelligence avant le langage. Paris: Flammarion, 1936. Pp. 292. 12 francs.

KLEIN, D. B. General psychology. New York: Holt, 1936. Pp. xiv+560. $2.50.

LOW, A. A. Studies in infant speech and thought. Part I. *Univ. Illinois Bull.,* 1936, **33**, No. 39. (*Illinois Med. & Dental Monog.,* 1936, **1**, No. 2.) Pp. 7+71. $1.00.

Manicomio, 1936, No. 2, 1-60. Mexico, D. F.: Sociedad para Estudios de Neurología y Psiquiatría.

MEAD, G. H. Movements of thought in the nineteenth century. Chicago: Univ. Chicago Press, 1936. Pp. xxix+519.

RUCKMICK, C. A. The psychology of feeling and emotion. New York & London: McGraw-Hill, 1936. Pp. xiii+529. $4.50.

SADLER, W. S. Theory and practice of psychiatry. St. Louis: Mosby, 1936. Pp. xxii+1231. $10.00.

SCHWARZ, O. L. Unconventional ethics. Washington, D. C.: Perennial Publications, 1936. Pp. 276. $3.50.

SEASHORE, C. E. Psychology of the vibrato in voice and instrument. *Univ. Iowa Stud.: Stud. Psychol. Music,* 1936, **3**. Pp. 159. Paper, $1.25; cloth, $1.75.

SKINNER, C. E. Educational psychology. New York: Prentice-Hall, 1936. Pp. xxvi+754. $3.20.

SORENSON, H. Statistics for students of psychology and education. New York & London: McGraw-Hill, 1936. Pp. viii+373. $3.50.

STUDEBAKER, J. W. Plain talk. Washington, D. C.: National Home Library Foundation, 1936. Pp. ix+166. $.25.

WAHL, M. Le mouvement dans la peinture. Paris: Alcan, 1936. Pp. viii+116. 15 fr.

WASHBURN, M. F. The animal mind. New York: Macmillan, 1936. Pp. xiii+536. $3.50.